Lower deviation probabilities for supercritical Galton–Watson processes

Klaus Fleischmann, Vitali Wachtel

Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, D-10117 Berlin, Germany

Abstract

There is a well-known sequence of constants c_n describing the growth of supercritical Galton–Watson processes Z_n. By lower deviation probabilities we refer to $\mathbf{P}(Z_n = k_n)$ with $k_n = \mathrm{o}(c_n)$ as n increases. We give a detailed picture of the asymptotic behavior of such lower deviation probabilities. This complements and corrects results known from the literature concerning special cases. Knowledge on lower deviation probabilities is needed to describe large deviations of the ratio Z_{n+1}/Z_n. The latter are important in statistical inference to estimate the offspring mean. For our proofs, we adapt the well-known Cramér method for proving large deviations of sums of independent variables to our needs.

Résumé

Les auteurs présentent une analyse détaillée des probabilités de déviations inférieures. Ces dernières sont nécessaires à la description du rapport Z_{n+1}/Z_n.

MSC: primary 60J80; secondary 60F10

Keywords: Supercritical Galton–Watson process; Local limit theorem; Large deviation; Cramér transform; Concentration function; Schröder equation; Böttcher equation

1. Introduction and statement of results

1.1. On the growth of supercritical processes

Let $Z = (Z_n)_{n\geqslant 0}$ denote a Galton–Watson process with offspring generating function

$$f(s) = \sum_{j\geqslant 0} p_j s^j, \quad 0 \leqslant s \leqslant 1, \tag{1}$$

☆ Supported by the German Science Foundation.

E-mail addresses: fleischm@wias-berlin.de (K. Fleischmann), vakhtel@wias-berlin.de (V. Wachtel).
URLs: http://www.wias-berlin.de/~fleischm (K. Fleischmann), http://www.wias-berlin.de/~vakhtel (V. Wachtel).

which is required to be non-degenerate, that is, $p_j < 1,\ j \geqslant 0$. Suppose that Z is supercritical, i.e. $f'(1) =: m \in (1, \infty)$. For simplicity, the initial state $Z_0 \geqslant 1$ is always assumed to be deterministic, and, if not noted otherwise, we set $Z_0 = 1$.

It is well-known (see, e.g., Asmussen and Hering [1, §3.5]) that

$$\text{there are } c_n > 0 \text{ such that a.s. } c_n^{-1} Z_n \underset{n\uparrow\infty}{\longrightarrow} \text{ some non-degenerate } W. \tag{2}$$

In this sense, the sequence of constants c_n describes the order of growth of Z. But $\mathbf{P}(W = 0)$ might be positive, more precisely, it equals the smallest root $q \in [0, 1)$ of $f(s) = s$, that is, it coincides with the extinction probability of Z. On the other hand, W restricted to $(0, \infty)$ has a (strictly) positive continuous density function denoted by w. Therefore the following *global limit theorem* holds:

$$\lim_{n\uparrow\infty} \mathbf{P}(Z_n \geqslant x c_n) = \int_x^\infty w(t)\,\mathrm{d}t, \quad x > 0. \tag{3}$$

The normalizing sequence $(c_n)_{n\geqslant 0}$ can be chosen to have the following additional properties:

$$c_0 = 1 \text{ and } c_n < c_{n+1} \leqslant m c_n, \quad n \geqslant 0, \tag{4a}$$

$$c_n = m^n L\big(m^n\big) \quad \text{with } L \text{ slowly varying at infinity}, \tag{4b}$$

$$\lim_{x\uparrow\infty} L(x) \text{ exists; it is positive if and only if } \mathbf{E} Z_1 \log Z_1 < \infty. \tag{4c}$$

Because of (4b,c), we may (and subsequently shall) take

$$c_n := m^n \quad \text{if } \mathbf{E} Z_1 \log Z_1 < \infty. \tag{5}$$

1.2. Asymptotic local behavior of Z, and main purpose

A local limit theorem related to (3) is due to Dubuc and Seneta [8], see also [1, §3.7]. To state it we need the following definition.

Definition 1 *(Type (d, μ)).* We say the offspring generating function f *is of type* (d, μ), if $d \geqslant 1$ is the greatest common divisor of the set $\{j - \ell\colon\ j \neq \ell,\ p_j p_\ell > 0\}$, and $\mu \geqslant 0$ is the minimal $j \geqslant 0$ for which $p_j > 0$.

Here is the announced *local limit theorem*. Suppose f is of type (d, μ). Take $x > 0$, and consider integers $k_n \geqslant 1$ such that $k_n / c_n \to x$ as $n \uparrow \infty$. Then, for each $j \geqslant 1$,

$$\lim_{n\uparrow\infty} \Big(c_n \mathbf{P}\{Z_n = k_n | Z_0 = j\} - d\,\mathbb{1}_{\{k_n \equiv j\mu^n (\operatorname{mod} d)\}} w_j(x) \Big) = 0, \tag{6}$$

where $w_j := \sum_{\ell=1}^{j} \binom{j}{\ell} q^{j-\ell} w^{*\ell}$, and $w^{*\ell}$ denotes the ℓ-fold convolution of the density function w. In particular, in our standard case $Z_0 = 1$ and if additionally $k_n \equiv \mu^n (\operatorname{mod} d)$, then

$$\mathbf{P}(Z_n = k_n) \sim d c_n^{-1} w(k_n / c_n) \quad \text{as } n \uparrow \infty \tag{7}$$

(with the usual meaning of the symbol $\sim$ as the ratio converges to 1).

Statement (6) [and especially (7)] can be considered as describing the local behavior of supercritical Galton–Watson processes in the region of *normal* deviations (from the growth of the c_n; 'deviations' are meant here in a multiplicative sense, related to the multiplicative nature of branching). But what about $\mathbf{P}(Z_n = k_n)$ when $k_n / c_n \to 0$ or ∞? In these cases we speak of *lower* and *upper* (local) deviation probabilities, respectively.

There are good reasons to be interested in the behavior of these probabilities. Lower deviations of Z_n are closely related to large deviations of Z_{n+1}/Z_n (see Ney and Vidyashankar [14, Section 2.3]). The latter are important in statistical inference for supercritical Galton–Watson processes, since Z_{n+1}/Z_n is the well-known Lotka–Nagaev estimator of the offspring mean.

The *main purpose* of the present paper is to study lower deviation probabilities in their own right and to provide a detailed picture (see Theorems 4 and 6 below). Applications of our results for large deviations of Z_{n+1}/Z_n can be found in [11].

Here is the program for the remaining part of the introduction. After introducing and discussing a basic dichotomy, we review in Section 1.4 what is known on lower deviations from the literature, before we state our results in Sections 1.5 and 1.6.

1.3. A dichotomy for supercritical processes

Recalling that f denotes the offspring generating function, q the extinction probability, and m the mean,

$$\text{set } \gamma := f'(q), \text{ and define } \alpha \text{ by } \gamma = m^{-\alpha}. \tag{8}$$

Note that $\gamma \in [0, 1)$ and $\alpha \in (0, \infty]$. We introduce the following notion, reflecting a crucial dichotomy for supercritical Galton–Watson processes.

Definition 2 *(Schröder and Böttcher case).* For our supercritical offspring law we distinguish between the *Schröder* and the *Böttcher* case, in dependence on whether $p_0 + p_1 > 0$ or $= 0$.

Obviously, f is of Schröder type if and only if $\gamma > 0$, if and only if $\alpha < \infty$.

Next we want to collect a few basic facts from the literature concerning that dichotomy. Clearly, f can be considered as a function on the closed unit disc D in the complex plane. As usual, denote by f_n the nth iterate of f.

We start with the *Schröder case*. Here it is well-known (see, e.g., [1, Lemma 3.7.2 and Corollary 3.7.3]) that

$$\mathsf{S}_n(z) := \frac{f_n(z) - q}{\gamma^n} \underset{n\uparrow\infty}{\longrightarrow} \text{some } \mathsf{S}(z) =: \sum_{j=0}^{\infty} \nu_j z^j, \quad z \in D. \tag{9}$$

Moreover, the convergence is uniform on each compact subset of the interior D° of D. Furthermore, the function S restricted to the reals is the unique solution of the so-called *Schröder functional equation* (see, e.g., Kuczma [13, Theorem 6.1, p. 137]),

$$\mathsf{S}\big(f(s)\big) = \gamma \mathsf{S}(s), \quad 0 \leqslant s \leqslant 1, \tag{10}$$

satisfying

$$\mathsf{S}(q) = 0 \quad \text{and} \quad \lim_{s\to q} \mathsf{S}'(s) = 1. \tag{11}$$

As a consequence of (9),

$$\lim_{n\uparrow\infty} \gamma^{-n} \mathbf{P}(Z_n = k) = \nu_k, \quad k \geqslant 1. \tag{12}$$

Consequently, in the Schröder case, these extreme (k is fixed) lower deviation probabilities $\mathbf{P}(Z_n = k)$ are positive and decay to 0 with order γ^n. On the other hand, the characteristic $\alpha \in (0, \infty)$ describes the behavior of the limiting quantities $w(x)$ and $\mathbf{P}(W \leqslant x)$ as $x \downarrow 0$. In fact, according to Biggins and Bingham [4], there is a continuous, positive multiplicatively periodic function V such that

$$x^{1-\alpha} w(x) = V(x) + \mathrm{o}(1) \quad \text{as } x \downarrow 0. \tag{13}$$

Dubuc [6] has shown that the function V can be replaced by a constant $V_0 > 0$ if and only if

$$\mathsf{S}\big(\varphi(h)\big) = K_0 h^{-\alpha}, \quad h \geqslant 0, \tag{14}$$

for some constant $K_0 > 0$, where $\varphi = \varphi_W$ denotes the Laplace transform of W,

$$\varphi_W(h) := \mathbf{E}\,\mathrm{e}^{-hW}, \quad h \geqslant 0. \tag{15}$$

We mention that condition (14) is certainly fulfilled if Z is embeddable (see [1, p. 96]) into a continuous-time Galton–Watson process (as in the case of a geometric offspring law, see Example 3 below).

Now we turn to the *Böttcher case*. Here $\mu \geqslant 2$ (recall Definition 1). Clearly, opposed to (12), extreme lower deviation probabilities disappear, even $\mathbf{P}(Z_n < \mu^n) = 0$ for all $n \geqslant 1$. Evidently,

$$\mathbf{P}\big(Z_n = \mu^n\big) = \mathbf{P}\big(Z_{n-1} = \mu^{n-1}\big)\, p_\mu^{(\mu^{n-1})}. \tag{16}$$

Hence,

$$\mathbf{P}\big(Z_n = \mu^n\big) = \prod_{j=0}^{n-1} p_\mu^{(\mu^j)} = \exp\left[\frac{\mu^n - 1}{\mu - 1}\log p_\mu\right]. \tag{17}$$

Next, $\mathbf{P}(Z_n = \mu^n + 1) = \mathbf{P}(Z_{n-1} = \mu^{n-1})\mu^{n-1} p_{\mu+1} p_\mu^{\mu^{n-1}-1}$. Thus, from (16),

$$\mathbf{P}\big(Z_n = \mu^n + 1\big) = p_\mu^{-1} p_{\mu+1} \mu^{n-1} \mathbf{P}\big(Z_n = \mu^n\big). \tag{18}$$

For simplification, consider for the moment the special case $p_{\mu+j} > 0,\ j \geqslant 0$. Then, as in (18), for fixed $k \geqslant 0$ and some positive constants C_k,

$$\mathbf{P}\big(Z_n = \mu^n + k\big) \sim C_k \mu^{nk} \mathbf{P}\big(Z_n = \mu^n\big) \quad \text{as } n \uparrow \infty. \tag{19}$$

Consequently, in contrast to (12) in the Schröder case, here the lower positive deviation probabilities $\mathbf{P}(Z_n = \mu^n + k)$ do *not* have a uniform order of decay. But by (19),

$$\mu^{-n} \log \mathbf{P}\big(Z_n = \mu^n + k\big) \underset{n\uparrow\infty}{\longrightarrow} \log p_\mu, \quad k \geqslant 0. \tag{20}$$

That is, on a *logarithmic* scale, we have again a uniform order, namely the order $-\mu^n$.

Turning back to the general Böttcher case,

$$\lim_{n\uparrow\infty} \big(f_n(s)\big)^{(\mu^{-n})} =: \mathsf{B}(s), \quad 0 \leqslant s \leqslant 1, \tag{21}$$

exists, is continuous, positive, and satisfies the *Böttcher functional equation*

$$\mathsf{B}\big(f(s)\big) = \mathsf{B}^\mu(s), \quad 0 \leqslant s \leqslant 1, \tag{22}$$

with boundary conditions

$$\mathsf{B}(0) = 0 \quad \text{and} \quad \mathsf{B}(1) = 1 \tag{23}$$

(see, e.g., Kuczma [13, Theorem 6.9, p. 145]).

Recalling that $\mu \geqslant 2$ in the Böttcher case, define $\beta \in (0, 1)$ by

$$\mu = m^\beta. \tag{24}$$

According to [4, Theorem 3], there exists a positive and multiplicatively periodic function V^* such that

$$-\log \mathbf{P}(W \leqslant x) = x^{-\beta/(1-\beta)} V^*(x) + \mathrm{o}\big(x^{-\beta(1-\beta)}\big) \quad \text{as } x \downarrow 0. \tag{25}$$

If additionally $\log \varphi_W(h) \sim -\kappa h^\beta$ as $h \uparrow \infty$ for some constant $\kappa > 0$, then by Bingham [5, formula (4)],

$$-\log \mathbf{P}(W \leqslant x) \sim \beta^{-1}(1-\beta)(\kappa\beta)^{1/(1-\beta)} x^{-\beta/(1-\beta)} \quad \text{as } x \downarrow 0. \tag{26}$$

1.4. Lower deviation probabilities in the literature

What else is known in the literature on lower deviation probabilities of Z? In the *Schröder case* $(0 < \alpha < \infty)$, Athreya and Ney [2] proved that in case of mesh $d = 1$ and $\mathbf{E}Z_1^2 < \infty$, for every $\varepsilon \in (0, \eta)$, where

$$\eta := m^{\alpha/(3+\alpha)} > 1, \tag{27}$$

there exists a positive constant C_ε such that for all $k \geqslant 1$,

$$\left| m^n \mathbf{P}(Z_n = k) - w\left(\frac{k}{m^n}\right)\right| \leqslant C_\varepsilon \frac{\eta^{-n}}{km^{-n}} + (\eta - \varepsilon)^{-n}. \tag{28}$$

The estimate (28) allows us to get some information on lower deviation probabilities. Indeed, in the general Schröder case, from (13),

$$w(x) \asymp x^{\alpha-1} \quad \text{as } x \downarrow 0 \tag{29}$$

(meaning that there are positive constants C_1 and C_2 such that $C_1 x^{\alpha-1} \leqslant w(x) \leqslant C_2 x^{\alpha-1}$, $0 < x \leqslant 1$). Together with (28) this implies

$$\mathbf{P}(Z_n = k_n) = m^{-n} w\big(k_n/m^n\big)\left[1 + \mathrm{O}\left(\frac{m^{\alpha n}}{k_n^{\alpha}\eta^n} + \frac{m^{(\alpha-1)n}}{k_n^{\alpha-1}(\eta-\varepsilon)^n}\right)\right] \quad \text{as } n \uparrow \infty. \tag{30}$$

We want to show that in important special cases the O-expression is actually an o(1). Recalling the definition (27) of η, one easily verifies that $m^{\alpha n}/k_n^{\alpha}\eta^n \to 0$ (as $n \uparrow \infty$) if and only if $k_n/m^{n(2+\alpha)/(3+\alpha)} \to \infty$. Concerning the second O-term, if additionally $\alpha \leqslant 1$, then $m^{(\alpha-1)n}/k_n^{\alpha-1} \leqslant 1$ provided that $k_n \leqslant m^n$. Hence, here $m^{(\alpha-1)n}/(k_n^{\alpha-1}(\eta-\varepsilon)^n)$ converges to zero if $\eta - \varepsilon > 1$. On the other hand, if $\alpha > 1$ and $k_n/m^{n(2+\alpha)/(3+\alpha)} \to \infty$ (which we needed for the first term), then $m^{(\alpha-1)n}/(k_n^{\alpha-1}(\eta-\varepsilon)^n) \to 0$ provided that additionally $\varepsilon \leqslant m^{\alpha/(3+\alpha)} - m^{(\alpha-1)/(3+\alpha)}$. Altogether, in the Schröder case and under the assumptions in [2],

$$\mathbf{P}(Z_n = k_n) = m^{-n} w\big(k_n/m^n\big)\big(1 + \mathrm{o}(1)\big) \quad \text{as } n \uparrow \infty \tag{31}$$

provided that both $k_n \leqslant m^n$ and $k_n/m^{n(2+\alpha)/(3+\alpha)} \to \infty$.

In [2] it is also mentioned that according to an unpublished manuscript of S. Karlin, in the Schröder case, for each embeddable process Z of finite second moment,

$$\lim_{n\uparrow\infty} \frac{m^{\alpha n}}{k_n^{\alpha-1}} \mathbf{P}(Z_n = k_n) \quad \text{exists in } (0,\infty), \text{ provided that } k_n = \mathrm{o}\big(m^n\big). \tag{32}$$

In the present situation, as we remarked after (13), $w(x) \sim V_0 x^{\alpha-1}$ as $x \downarrow 0$ with $V_0 > 0$. Hence, from (32), for some constant $C > 0$,

$$\mathbf{P}(Z_n = k_n) \sim C m^{-n} w\big(k_n/m^n\big) \quad \text{as } n \uparrow \infty, \tag{33}$$

which is compatible with (31).

Intuitively, the asymptotic behavior of lower deviation probabilities should be more related to characteristics such as α and β than to the tail of the offspring distribution. Thus one can expect that it is possible to describe lower deviation probabilities successfully without the second moment assumption used in [2]. Actually, in [14, Theorem 1] one finds the following *claim*.

Suppose $p_0 = 0$ and $\mathbf{E}Z_1 \log Z_1 < \infty$. Then there exist positive constants $C_1 < C_2$ such that for $k_n \to \infty$ with $k_n = \mathrm{O}(m^n)$ as $n \uparrow \infty$,

$$C_1 \leqslant \liminf_{n\uparrow\infty} \frac{\mathbf{P}(Z_n = k_n)}{A_n} \leqslant \limsup_{n\uparrow\infty} \frac{\mathbf{P}(Z_n = k_n)}{A_n} \leqslant C_2, \tag{34}$$

where

$$A_n := \begin{cases} p_1^n k_n^{\alpha-1} & \text{if } \alpha < 1, \\ \theta_n p_1^n & \text{if } \alpha = 1, \\ m^{-n} & \text{if } 1 < \alpha \leqslant \infty, \end{cases} \tag{35}$$

and $\theta_n := [n + 1 - \log k_n/\log m]$. Furthermore, if $k_n = m^{n-\ell_n}$ for natural numbers $\ell_n = \mathrm{O}(n)$ as $n \uparrow \infty$, then

$$\lim_{n\uparrow\infty} A_n^{-1}\mathbf{P}(Z_n = k_n) =: C_{\lim} \text{ exists in } (0,\infty). \tag{36}$$

Unfortunately, this claim is not true as it stands. In fact, consider first the following example.

Example 3 *(Geometric offspring law).* Consider the offspring generating function

$$f(s) = \frac{s}{m - (m-1)s} = \sum_{j=1}^{\infty} m^{-1}\big(1 - m^{-1}\big)^{j-1} s^j, \quad 0 \leqslant s \leqslant 1, \tag{37}$$

(with mean $m > 1$). Obviously, here $q = 0$, $\gamma = m^{-1}$, hence $\alpha = 1$. For the nth iterate one easily gets

$$f_n(s) = \frac{s}{m^n - (m^n - 1)s} = \sum_{j=1}^{\infty} m^{-n}\big(1 - m^{-n}\big)^{j-1} s^j. \tag{38}$$

Thus,

$$\mathbf{P}(Z_n = k) = m^{-n}\bigl(1 - m^{-n}\bigr)^{k-1} \leqslant m^{-n}, \tag{39}$$

for all $n, k \geqslant 1$. On the other hand, since $p_1 = m^{-1}$, by claim (34) there is a constant $C > 0$ such that for the considered k_n,

$$\mathbf{P}(Z_n = k_n) \geqslant C\theta_n m^{-n} \tag{40}$$

for n large enough. If, for example, $k_n = m^{n/2}$ then $\theta_n \to \infty$, and (40) contradicts (39).

Consequently, the left-hand part of claim (34) cannot be true in the case $\alpha = 1$. Next, in the case $1 < \alpha < \infty$, we compare the claim with (31). In fact, under the assumptions in [2], if additionally $k_n = \mathrm{o}(m^n)$ but

$$\frac{k_n}{m^{n(2+\alpha)/(3+\alpha)}} \to \infty \quad \text{as } n \uparrow \infty,$$

then by (31) and (29),

$$\mathbf{P}(Z_n = k_n) \asymp m^{-n}\left(\frac{k_n}{m^n}\right)^{\alpha-1}. \tag{41}$$

Thus, we get $\mathbf{P}(Z_n = k_n) = \mathrm{o}(m^{-n})$ which contradicts the positivity of $C_{\lim}$ in claim (36), hence of C_1 in claim (34). Finally, in the case $\alpha = \infty$, the proof of Lemma 5 in [14] is incorrect. In fact, the statement (82) there is wrong. Right calculations instead lead to $C_1 = 0$ in this case.

Summarizing, for each value of $\alpha \in [1, \infty]$, the claimed positivity of C_1 in (34) is not always true. (Some more discussion on the claim (34) can be found in our original preprint [10, Section 1.5].)

Actually, it is wrong to distinguish between velocity cases as in (35). The only needed velocity case differentiation is the mentioned dichotomy of Definition 2. This we will explain in the next two sections. Moreover, there we also remove the $Z_1 \log Z_1$-moment assumption, used in [14].

1.5. Lower deviations in the Schröder case

We start by stating our results on lower deviation probabilities in the Schröder case. Recall that here $\mu = 0$ or 1.

Theorem 4 *(Schröder case). Let the offspring law be of the Schröder type and of type (d, μ). Then*

$$\sup_{k \geqslant \tilde{k} \text{ with } k \equiv \mu (\mathrm{mod}\, d),\ j \geqslant 0} \left| \frac{m^j c_{a_k}}{d w(k/(m^j c_{a_k}))} \mathbf{P}(Z_{a_k + j} = k) - 1 \right| \underset{\tilde{k} \uparrow \infty}{\longrightarrow} 0 \tag{42}$$

and

$$\sup_{k \geqslant \tilde{k},\ j \geqslant 0} \left| \frac{\mathbf{P}(0 < Z_{a_k + j} \leqslant k)}{\mathbf{P}(0 < W < k/(m^j c_{a_k}))} - 1 \right| \underset{\tilde{k} \uparrow \infty}{\longrightarrow} 0, \tag{43}$$

where for $k \geqslant 1$ fixed, we put $a_k := \min\{\ell \geqslant 1\colon\ c_\ell \geqslant k\}$.

It seems to be convenient to expose the following immediate implication.

Corollary 5 *(Schröder case). Under the conditions of Theorem* 4, *for $k_n \leqslant c_n$ satisfying $k_n \to \infty$, we have*

$$\sup_{k \in [k_n, c_n] \text{ with } k \equiv \mu (\mathrm{mod}\, d)} \left| \frac{m^{n - a_k} c_{a_k}}{d w(k/(m^{n-a_k} c_{a_k}))} \mathbf{P}(Z_n = k) - 1 \right| \underset{n \uparrow \infty}{\longrightarrow} 0 \tag{44}$$

and

$$\sup_{k \in [k_n, c_n]} \left| \frac{\mathbf{P}(0 < Z_n \leqslant k)}{\mathbf{P}(0 < W < k/(m^{n-a_k} c_{a_k}))} - 1 \right| \underset{n \uparrow \infty}{\longrightarrow} 0. \tag{45}$$

The appearing of the a_k in the theorem and corollary, depending on k and on the sequence of the c_n, looks a bit disturbing, so we have to discuss it. First assume additionally that $\mathbf{E}Z_1 \log Z_1 < \infty$. Since here we set $c_n = m^n$ [recall (5)], from (44) we obtain the a_k-free formula

$$\mathbf{P}(Z_n = k) = dm^{-n} w\big(k/m^n\big)\big(1 + \mathrm{o}(1)\big). \tag{46}$$

Also, comparing this with (7), we see that under this $Z_1 \log Z_1$-moment condition in the Schröder case, $m^{-n} w(k/m^n)$ describes not only normal deviation probabilities but also lower ones.

On the other hand, without this additional moment condition, recalling property (4b), we have $c_n = m^n L(m^n)$ with L slowly varying at infinity. Hence,

$$\frac{1}{m^{n-a_k} c_{a_k}} = \frac{1}{c_n} \frac{L(m^n)}{L(m^{a_k})}, \quad \text{thus } \frac{k}{c_{a_k} m^{n-a_k}} = \frac{k}{c_n} \frac{L(m^n)}{L(m^{a_k})}. \tag{47}$$

Therefore, from (44),

$$\frac{c_n \mathbf{P}(Z_n = k)}{dw(k/c_n)} = \frac{L(m^n)}{L(m^{a_k})} \frac{w(kL(m^n)/c_n L(m^{a_k}))}{w(k/c_n)} \big(1 + \mathrm{o}(1)\big). \tag{48}$$

Using now (13), we find

$$\frac{c_n \mathbf{P}(Z_n = k)}{dw(k/c_n)} = \left(\frac{L(m^n)}{L(m^{a_k})}\right)^{\alpha} \frac{V(kL(m^n)/c_n L(m^{a_k}))}{V(k/c_n)} \big(1 + \mathrm{o}(1)\big). \tag{49}$$

Next we want to expel the disturbing a_k from this formula.

It is well known (Seneta [16, p. 23]) that the regularly varying function $x \mapsto xL(x)$ asymptotically equals a (strictly) increasing, continuous, regularly varying function $x \mapsto R(x) := xL_1(x)$ with slowly varying L_1. Hence, $L(x) \sim L_1(x)$ as $x \uparrow \infty$. Using now [16, Lemma 1.3], we conclude that the inverse function R^* of R equals $x \mapsto xL^*(x)$, where L^* is again a slowly varying function.

Put $x_k := R^*(k)$. Then $k = x_k L_1(x_k)$ by the definition of R^*. Recalling that $x_k = kL^*(k)$, we get the identity

$$L^*(k) L_1(x_k) = 1, \quad k \geqslant 1. \tag{50}$$

For n fixed, define $b_k := \min\{\ell \geqslant 1\colon m^\ell L_1(m^\ell) \geqslant k\}$. Combined with $x_k L_1(x_k) = k$ we get

$$m^{b_k} L_1\big(m^{b_k}\big) \geqslant x_k L_1(x_k) > m^{b_k - 1} L_1\big(m^{b_k - 1}\big). \tag{51}$$

But $x \mapsto xL_1(x)$ is increasing, and the previous chain of inequalities immediately gives

$$m^{b_k} \geqslant x_k > m^{b_k - 1}. \tag{52}$$

By (4b),

$$c_{b_k+1} = m^{b_k+1} L\big(m^{b_k+1}\big) = m \frac{L(m^{b_k+1})}{L_1(m^{b_k})} m^{b_k} L_1\big(m^{b_k}\big) \geqslant k \tag{53}$$

for all n sufficiently large. Here, in the last step we used $m > 1$, that the slowly varying functions L and L_1 are asymptotically equivalent, and the definition of b_k. Now $c_{b_k+1} \geqslant k$ implies

$$b_k + 1 \geqslant a_k, \tag{54}$$

by the definition of a_k. On the other hand,

$$m^{a_k+1} L_1\big(m^{a_k+1}\big) = m \frac{L_1(m^{a_k+1})}{L(m^{a_k})} c_{a_k} \geqslant k \tag{55}$$

for all n sufficiently large. Here, in the last step we used the definition of a_k. This gives

$$a_k + 1 \geqslant b_k, \tag{56}$$

by the definition of b_k. Entering with (56) and (54) into (52), we get

$$m^{a_k+1} \geqslant x_k > m^{a_k - 2} \quad \text{for all } k \text{ sufficiently large.} \tag{57}$$

Therefore, recalling (50),

$$L(m^{a_k}) \sim L(x_k) \sim L_1(x_k) \sim \frac{1}{L^*(k)} \quad \text{as } k \uparrow \infty. \tag{58}$$

Entering this into (49) gives

$$\frac{c_n \mathbf{P}(Z_n = k)}{d w(k/c_n)} = \big[L\big(m^n\big) L^*(k)\big]^\alpha \frac{V(kL(m^n)L^*(k)/c_n)}{V(k/c_n)}\big(1 + \mathrm{o}(1)\big), \tag{59}$$

which contains L^* instead of the a_k.

Note also that such reformulation of (44) reminds one of the classical Cramér theorem (see, for example, Petrov [15, §VIII.2]) on large deviations for sums of independent random variables. There the ratio of a tail probability of a sum of independent variables and the corresponding normal law expression is considered. The crucial role in Cramér's theorem is played by the so-called Cramér series $\lambda(s) := \sum_{k=0}^{\infty} \lambda_k s^k$, where the coefficients λ_k depend on the cumulants of the summands. For the lower deviation probabilities of supercritical Galton–Watson processes we have a more complex situation: It is not at all clear, how to find the input data L, L^*, V [entering into (59)] based only on the knowledge of the offspring generating function f.

It was already noted after (13) that if Z is embeddable into a continuous-time Galton–Watson process then $V(x) \equiv V_0$. Consequently, for embeddable processes, (59) takes the slightly simpler form

$$\frac{c_n \mathbf{P}(Z_n = k)}{d w(k/c_n)} = \big[L\big(m^n\big) L^*(k)\big]^\alpha \big(1 + \mathrm{o}(1)\big). \tag{60}$$

On the other hand, if V is not constant, the ratio $V(kL(m^n)L^*(k)/c_n)/V(k/c_n)$ gives oscillations in the asymptotic behavior of $c_n \mathbf{P}(Z_n = k)/w(k/c_n)$. But the influence of the function V is relatively small. Indeed, from the continuity and multiplicative periodicity of $V(x)$ we see that $0 < V_1 \leqslant V(x) \leqslant V_2 < \infty$, $x > 0$, for some constants V_1, V_2. Therefore, the oscillations are in the interval $[V_1/V_2, V_2/V_1]$, that is, from (59) we obtain

$$\frac{V_1}{V_2}\big[L\big(m^n\big) L^*(k)\big]^\alpha \big(1 + \mathrm{o}(1)\big) \leqslant \frac{c_n \mathbf{P}(Z_n = k)}{d w(k/c_n)} \leqslant \frac{V_2}{V_1}\big[L\big(m^n\big) L^*(k)\big]^\alpha \big(1 + \mathrm{o}(1)\big). \tag{61}$$

Note also that for many offspring distributions the bounds V_1 and V_2 may be chosen close to each other. This "near-constancy" phenomenon was studied by Dubuc [7] and by Biggins and Bingham [3,4].

1.6. Lower deviations in the Böttcher case

Recall that $\mu \geqslant 2$ in the Böttcher case.

Theorem 6 *(Böttcher case). Let the offspring law be of the Böttcher type and of type (d, μ). Then there exist positive constants B_1 and B_2 such that for all $k_n \equiv \mu^n \,(\mathrm{mod}\, d)$ with $k_n \geqslant \mu^n$ but $k_n = \mathrm{o}(c_n)$,*

$$-B_1 \leqslant \liminf_{n \uparrow \infty} \mu^{b_n - n} \log\big[c_n \mathbf{P}(Z_n = k_n)\big] \tag{62a}$$

$$\leqslant \limsup_{n \uparrow \infty} \mu^{b_n - n} \log\big[c_n \mathbf{P}(Z_n = k_n)\big] \leqslant -B_2, \tag{62b}$$

where $b_n := \min\{\ell\colon\ c_\ell \mu^{n-\ell} \geqslant 2k_n\}$. The inequalities remain true if one replaces $c_n \mathbf{P}(Z_n = k_n)$ by $\mathbf{P}(Z_n \leqslant k_n)$, where in this integral case the assumption $k_n \equiv \mu_n \ (\mathrm{mod}\, d)$ is superflous.

Let us add at this place the following remark.

Remark 7 *(Behavior of w at* 0*).* In analogy with (29), in the Böttcher case one has

$$\log w(x) \asymp -x^{-\beta/(1-\beta)} \quad \text{as } x \downarrow 0 \tag{63}$$

with β from (24). This can be shown using techniques from the proof of Theorem 6; see [10, Remark 16].

Our results in the Böttcher case are much weaker than the results in the Schröder case: We got only logarithmic bounds. But this is not unexpected, recall our discussion around (20).

Repeating arguments as we used to obtain (59), from Theorem 6 we get

$$\frac{\log[c_n\mathbf{P}(Z_n=k_n)]}{(k_n/c_n)^{-\beta/(1-\beta)}} \asymp -\big[L^*\big(k_n/m^{\beta n}\big)L^{1/(1-\beta)}\big(m^n\big)\big]^\beta \quad \text{as } n\uparrow\infty, \tag{64}$$

where L^* is such that $R_1(x):=x^{(1-\beta)}L(x)$ and $R_2(x):=x^{1/(1-\beta)}L^*(x)$ are asymptotic inverses, i.e. $R_1(R_2(x))\sim x$ and $R_2(R_1(x))\sim x$ as $x\uparrow\infty$.

Taking into account (63), we conclude that

$$\frac{\log[c_n\mathbf{P}(Z_n=k_n)]}{\log w(k_n/c_n)} \asymp \big[L^*\big(k_n/m^{\beta n}\big)L^{1/(1-\beta)}\big(m^n\big)\big]^\beta \quad \text{as } n\uparrow\infty. \tag{65}$$

2. Cramér transforms applied to Galton–Watson processes

Our way to prove Theorems 4 and 6 is based on the well-known Cramér method (see, e.g., [15, Chapter 8]), which was developed to study large deviations for sums of independent random variables. A key in this method is the so-called *Cramér transform* defined as follows. A random variable $X(h)$ is called a Cramér transform (with parameter $h\in\mathbb{R}$) of the random real variable X if

$$\mathbf{E}\,\mathrm{e}^{\mathrm{i}tX(h)} = \frac{\mathbf{E}\,\mathrm{e}^{(h+\mathrm{i}t)X}}{\mathbf{E}\,\mathrm{e}^{hX}}, \quad t\in\mathbb{R}. \tag{66}$$

Of course, this transformation is well-defined if $\mathbf{E}\,\mathrm{e}^{hX}<\infty$.

In what follows, we will *always assume* that our offspring law additionally satisfies $p_0=0$. This condition is not crucial but allows a slightly simplified exposition of auxiliary results formulated in Lemma 12 below and of the proof of Theorem 4 in Section 3.1 (see also Remark 16 below).

2.1. Basic estimates

Fix an offspring law of type (d,μ). Let $n\geqslant 1$. Since $Z_n>0$, the Cramér transforms $Z_n(-h/c_n)$ exist for all $h\geqslant 0$. Clearly, $\mathbf{E}\,\mathrm{e}^{\mathrm{i}tZ_n(-h/c_n)}=f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t})/f_n(\mathrm{e}^{-h/c_n})$. We want to derive upper bounds for $f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t})$ on $\{t\in\mathbb{R}\colon c_n^{-1}\pi d^{-1}\leqslant|t|\leqslant\pi d^{-1}\}$. For this purpose, it is convenient to decompose the latter set into $\bigcup_{j=1}^n J_j$ where

$$J_j:=\big\{t\colon c_j^{-1}\pi d^{-1}\leqslant|t|\leqslant c_{j-1}^{-1}\pi d^{-1}\big\}, \quad j\geqslant 1. \tag{67}$$

To prepare for this, we start with the following generalization of [8, Lemma 2].

Lemma 8 *(Preparation). Fix $\varepsilon\in(0,1)$. There exists $\theta=\theta(\varepsilon)\in(0,1)$ such that*

$$\big|f_\ell\big(\mathrm{e}^{-h/c_\ell+\mathrm{i}t/c_\ell}\big)\big|\leqslant\theta, \quad \ell\geqslant 0,\ h\geqslant 0,\ t\in J_\varepsilon:=\big\{t\colon \varepsilon\pi d^{-1}\leqslant|t|\leqslant\pi d^{-1}\big\}.$$

Proof. Put $g_{h,t}(x):=\mathrm{e}^{-hx+\mathrm{i}tx}$, $h,x\geqslant 0$, $t\in\mathbb{R}$. Evidently,

$$\begin{aligned}\big|g_{h,t}(x)-g_{h,t}(y)\big| &= \big|\mathrm{e}^{-hx}\big(\mathrm{e}^{\mathrm{i}tx}-\mathrm{e}^{\mathrm{i}ty}\big)+\mathrm{e}^{\mathrm{i}ty}\big(\mathrm{e}^{-hx}-\mathrm{e}^{-hy}\big)\big| \\ &\leqslant \big|\mathrm{e}^{\mathrm{i}tx}-\mathrm{e}^{\mathrm{i}ty}\big|+\big|\mathrm{e}^{-hx}-\mathrm{e}^{-hy}\big| \leqslant \big(h+|t|\big)|x-y|.\end{aligned} \tag{68}$$

This means that for $H\geqslant 1$ and $T\geqslant\pi d^{-1}$ fixed, $\mathcal{G}:=\{g_{h,t};\,0\leqslant h\leqslant H,\,|t|\leqslant T\}$ is a family of uniformly bounded and equi-continuous functions on $\mathbb{R}_+$. Therefore, by (2),

$$f_\ell(\mathrm{e}^{-h/c_\ell+\mathrm{i}t/c_\ell})=\mathbf{E}g_{h,t}(Z_\ell/c_\ell)\to\mathbf{E}g_{h,t}(W) \quad \text{as } \ell\uparrow\infty, \tag{69}$$

uniformly on $\mathcal{G}$ (see, e.g., Feller [9, Corollary in Chapter VIII, §1, p. 252]). Since $W>0$ has an absolutely continuous distribution, and $t\in J_\varepsilon$ implies $|t|\leqslant T$,

$$\sup_{0\leqslant h\leqslant H,\ t\in J_\varepsilon}\big|\mathbf{E}\,\mathrm{e}^{-hW+\mathrm{i}tW}\big|<1. \tag{70}$$

From (69) and (70) it follows that there exist $\delta_1 \in (0,1)$ and ℓ_0 such that

$$\sup_{0\leqslant h\leqslant H,\ t\in J_\varepsilon} \left| f_\ell\left(\mathrm{e}^{-h/c_\ell+\mathrm{i}t/c_\ell}\right)\right| \leqslant \delta_1, \quad \ell > \ell_0. \tag{71}$$

On the other hand, $\bigcup_{\ell=0}^{\ell_0}\{\mathrm{e}^{-h/c_\ell+\mathrm{i}t/c_\ell};\ h\geqslant 0,\ t\in J_\varepsilon\}$ is a subset of a compact subset K of the unit disc D, where K does not contain the dth roots of unity. Thus for some $\delta_2 \in (0,1)$,

$$\sup_{0\leqslant h\leqslant H,\ t\in J_\varepsilon} \left| f_\ell\left(\mathrm{e}^{-h/c_\ell+\mathrm{i}t/c_\ell}\right)\right| \leqslant \delta_2, \quad \ell \leqslant \ell_0. \tag{72}$$

In fact, from Definition 1,

$$f_\ell(z) = \sum_{j=0}^{\infty} \mathbf{P}\left(Z_\ell = \mu^\ell + jd\right) z^{\mu^\ell + jd}, \quad \ell \geqslant 0,\ z \in D, \tag{73}$$

implying

$$\left|f_\ell(z)\right| \leqslant \left|\sum_{j=0}^{\infty} \mathbf{P}\left(Z_\ell = \mu^\ell + jd\right) z^{jd}\right|. \tag{74}$$

But the latter sum equals 1 if and only if z is a dth root of unity, that is, if it is of the form $\mathrm{e}^{2\pi\mathrm{i}/d}$.

Combining (71) and (72) gives the claim in the lemma under the addition that $h \leqslant H$. Consider now any $h > H$. In this case

$$\left| f_\ell\left(\mathrm{e}^{-h/c_\ell+\mathrm{i}t/c_\ell}\right)\right| \leqslant f_\ell\left(\mathrm{e}^{-1/c_\ell}\right). \tag{75}$$

By (2) we have

$$f_\ell\left(\mathrm{e}^{-h/c_\ell}\right) = \mathbf{E}\,\mathrm{e}^{-hZ_\ell/c_\ell} \to \mathbf{E}\,\mathrm{e}^{-hW} \in (0,1] \quad \text{as } \ell \uparrow \infty, \tag{76}$$

uniformly for h in compact subsets of $\mathbb{R}_+$. In particular,

$$\sup_{\ell\geqslant 1} f_\ell\left(\mathrm{e}^{-1/c_\ell}\right) < 1. \tag{77}$$

This completes the proof. □

The following lemma generalizes [8, Lemma 3].

Lemma 9 *(Estimates on $J_1, \ldots J_n$). There are constants $A > 0$ and $\theta \in (0,1)$ such that for $h \geqslant 0$, $t \in J_j$, and $1 \leqslant j \leqslant n$,*

$$\left| f_n\left(\mathrm{e}^{-h/c_n+\mathrm{i}t}\right)\right| \leqslant \begin{cases} A p_1^{n-j+1} & \text{in the Schröder case,} \\ \theta^{(\mu^{n-j+1})} & \text{in all cases.} \end{cases} \tag{78}$$

Proof. By (4a), we have $\varepsilon := \inf_{\ell\geqslant 1} c_{\ell-1}/c_\ell \in (0,1)$. If $t \in J_j$, $j \geqslant 1$, then evidently,

$$\pi d^{-1} \geqslant c_{j-1}|t| \geqslant c_{j-1}c_j^{-1}\pi d^{-1} \geqslant \varepsilon\pi d^{-1}, \tag{79}$$

hence $c_{j-1}t \in J_\varepsilon$. Thus, by Lemma 8,

$$U := \bigcup_{j=1}^{\infty}\left\{ f_{j-1}\left(\mathrm{e}^{-h+\mathrm{i}t}\right);\ h \geqslant 0,\ t \in J_j\right\} \subseteq \theta D \quad \text{with } 0 < \theta < 1. \tag{80}$$

From the representation (73), $f_\ell(z) \leqslant |z|^{(\mu^\ell)}$ for all $\ell \geqslant 0$ and $|z| \leqslant 1$. Hence, for all $z \in U \subseteq \theta D$ we have the bound $|f_\ell(z)| \leqslant \theta^{(\mu^\ell)}$. Thus, for $h \geqslant 0$, $t \in J_j$, and $1 \leqslant j \leqslant n$,

$$\left| f_n\left(\mathrm{e}^{-h/c_n+\mathrm{i}t}\right)\right| \leqslant f_{n-j+1}\left(\left|f_{j-1}\left(\mathrm{e}^{-h/c_n+\mathrm{i}t}\right)\right|\right) \leqslant \theta^{(\mu^{n-j+1})}, \tag{81}$$

which is the second claim in (78).

If additionally $p_1 > 0$, then by (9) (and our assumption $p_0 = 0$) we have that $p_1^{-\ell} f_\ell(z)$ converges as $\ell \uparrow \infty$, uniformly on each compact $K \subset D^\circ$. Therefore, there exists a constant $C = C(K)$ such that

$$\left|f_\ell(z)\right| \leqslant C p_1^\ell, \quad \ell \geqslant 0, \quad z \in K. \tag{82}$$

Consequently, iterating as in (81),

$$\left|f_n\left(\mathrm{e}^{-h/c_n+\mathrm{i}t}\right)\right| \leqslant C p_1^{n-j+1}, \quad h \geqslant 0,\ t \in J_j,\ 1 \leqslant j \leqslant n, \tag{83}$$

finishing the proof. □

2.2. On concentration functions

Fix for the moment $h \geqslant 0$ and $n \geqslant 1$. Denote by $\{X_j(h,n)\}_{j\geqslant 1}$ a sequence of independent random variables which equal in law the Cramér transform $Z_n(-h/c_n)$, that is

$$\mathbf{P}\big(X_1(h,n) = k\big) = \frac{\mathrm{e}^{-kh/c_n}}{f_n(\mathrm{e}^{-h/c_n})} \mathbf{P}(Z_n = k), \quad k \geqslant 1. \tag{84}$$

Put

$$S_\ell(h,n) := \sum_{j=1}^{\ell} X_j(h,n), \quad \ell \geqslant 1. \tag{85}$$

Note that

$$\mathbf{E}\,\mathrm{e}^{\mathrm{i}t S_\ell(h,n)} = \left(f_n\left(\mathrm{e}^{-h/c_n+\mathrm{i}t}\right)/f_n\left(\mathrm{e}^{-h/c_n}\right)\right)^\ell. \tag{86}$$

Recall the notation $\alpha \in (0,\infty]$ from (8).

Lemma 10 *(A concentration function estimate). For every $h \geqslant 0$, there is a constant $A(h)$ such that*

$$\sup_{n,k\geqslant 1} c_n \mathbf{P}\big(S_\ell(h,n) = k\big) \leqslant \frac{A(h)}{\ell^{1/2}}, \quad \ell \geqslant \ell_0 := 1 + \left[\frac{1}{\alpha}\right]. \tag{87}$$

Proof. It is known (see, for example, [15, Lemma III.3, p. 38]) that for arbitrary (real-valued) random variables X and every $\lambda, T > 0$,

$$Q(X;\lambda) := \sup_y \mathbf{P}(y \leqslant X \leqslant y + \lambda) \leqslant \left(\frac{96}{95}\right)^2 \max\left(\lambda, T^{-1}\right) \int_{-T}^{T} \left|\psi_X(t)\right| \mathrm{d}t \tag{88}$$

(with ψ_X the characteristic function of X). Applying this inequality to $X = S_{\ell_0}(h,n)$ with $T = \pi d^{-1}$ and $\lambda = 1/2$, using (86) we have

$$\sup_{k\geqslant 1} \mathbf{P}\big(S_{\ell_0}(h,n) = k\big) \leqslant C \int_{-\pi d^{-1}}^{\pi d^{-1}} \frac{|f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t})|^{\ell_0}}{f_n^{\ell_0}(\mathrm{e}^{-h/c_n})} \,\mathrm{d}t \tag{89}$$

for some constant C independent of h, n. By (76), for h fixed, $f_n(\mathrm{e}^{-h/c_n})$ is bounded away from zero, and consequently, there is a positive constant $C(h)$ such that

$$\sup_{k\geqslant 1} \mathbf{P}\big(S_{\ell_0}(h,n) = k\big) \leqslant C(h) \int_{-\pi d^{-1}}^{\pi d^{-1}} \left|f_n\left(\mathrm{e}^{-h/c_n+\mathrm{i}t}\right)\right|^{\ell_0} \mathrm{d}t. \tag{90}$$

Fist assume that $\alpha < \infty$ (Schröder case). Using the first inequality in (78), we get for $1 \leqslant j \leqslant n$,

$$\int_{J_j} \left|f_n\left(\mathrm{e}^{-h/c_n+\mathrm{i}t}\right)\right|^{\ell_0} \mathrm{d}t \leqslant A^{\ell_0} p_1^{(n-j+1)\ell_0} |J_j| \leqslant 2\pi d^{-1} A^{\ell_0} p_1^{(n-j+1)\ell_0} c_{j-1}^{-1}. \tag{91}$$

On the other hand,

$$\int_{-\pi d^{-1}/c_n}^{\pi d^{-1}/c_n} \left|f_n\left(\mathrm{e}^{-h/c_n+\mathrm{i}t}\right)\right|^{\ell_0} \mathrm{d}t \leqslant \frac{2\pi d^{-1}}{c_n}. \tag{92}$$

From (91) and (92), for some constant C,

$$c_n \int_{-\pi d^{-1}}^{\pi d^{-1}} \left|f_n\left(\mathrm{e}^{-h/c_n+\mathrm{i}t}\right)\right|^{\ell_0} \mathrm{d}t \leqslant C\left(1+\sum_{j=1}^{n} p_1^{(n-j+1)\ell_0} c_n c_{j-1}^{-1}\right). \tag{93}$$

But by (4a),

$$c_n \leqslant m^{n-j+1} c_{j-1}, \quad 1 \leqslant j \leqslant n. \tag{94}$$

Also, by the definition of ℓ_0 in (87) and α in (8), $p_1^{\ell_0} m = p_1^{1+[1/\alpha]-1/\alpha} < 1$. Hence the right-hand side of (93) is bounded in n. Thus, from (90) it follows that

$$\sup_{n,k\geqslant 1} c_n \mathbf{P}\left(S_{\ell_0}(h,n)=k\right) \leqslant C(h). \tag{95}$$

This estimate actually holds also in the Böttcher case, where $\ell_0 = 1$. Indeed, proceeding in the same way but using the second inequality in (78) instead, the sum expression in (93) has to be replaced by

$$\sum_{j=1}^{n} \theta^{(\mu^{n-j+1})} c_n c_{j-1}^{-1} \leqslant \sum_{j=1}^{n} \theta^{(\mu^{n-j+1})} m^{n-j+1} = \sum_{j=1}^{n} \theta^{(\mu^{j})} m^{j}, \tag{96}$$

which again is bounded in n.

Note that (95) is (87) restricted to $\ell = \ell_0$. Hence, from now on we may focus our attention to $\ell > \ell_0$. Let $Y_1, \ldots, Y_j$ be independent identically distributed random variables. Then by Kesten's inequality (see, e.g., [15, p. 57]), there is a constant C such that for $0 < \lambda' < 2\lambda$ the concentration function inequality

$$Q(Y_1 + \cdots + Y_j; \lambda) \leqslant \frac{C\lambda}{\lambda' j^{1/2}} Q(Y_1; \lambda)\left[1 - Q(Y_1; \lambda')\right]^{-1/2} \tag{97}$$

holds. We specialize to $Y_1 = S_{\ell_0}(h,n)$ and $\lambda' = \lambda = 1/2$. Note that $Q(Y_1; 1/2) = \sup_{k\geqslant 1} \mathbf{P}(S_{\ell_0}(h,n)=k) < 1$ in this case, since the random variable $X_1(h,n)$ is non-degenerate. But also as $n \uparrow \infty$ this quantity is bounded away from 1, which follows from (95). Consequently, $\inf_{n\geqslant 1}[1 - Q(Y_1; 1/2)] > 0$. Thus, using again (95), we infer

$$\sup_{n,k\geqslant 1} \mathbf{P}\left(S_{j\ell_0}(h,n)=k\right) \leqslant \frac{C_1(h)}{j^{1/2}} = \frac{C_2(h)}{(j\ell_0)^{1/2}}, \quad j \geqslant 1, \tag{98}$$

for some positive constants $C_1(h)$ and $C_2(h)$. If X and Y are independent random variables, then, $Q(X+Y; \lambda) \leqslant Q(X; \lambda)$ (see [15, Lemma III.1]). Thus for every $\ell > \ell_0$ we have the inequality

$$\sup_{n,k\geqslant 1} c_n \mathbf{P}\left(S_\ell(h,n)=k\right) \leqslant \sup_{n,k\geqslant 1} c_n \mathbf{P}\left(S_{[\ell/\ell_0]\ell_0}(h,n)=k\right). \tag{99}$$

Combining this bound once more with (98), the proof is finished. □

Remark 11 *(Special case $h = 0$).* Note that $S_\ell(0,n)$ equals in law to Z_n conditioned to $Z_0 = \ell$. Therefore, by Lemma 10,

$$\sup_{k\geqslant 1} \mathbf{P}(Z_n = k | Z_0 = \ell) \leqslant \frac{A(0)}{\ell^{1/2} c_n}, \quad n \geqslant 1,\ \ell \geqslant \ell_0. \tag{100}$$

In particular, if $\alpha > 1$, implying $\ell_0 = 1$, in (100) all initial states Z_0 are possible. In the special case $Z_0 = 1$, inequality (100) generalizes the upper estimate in [14, (10)] to processes without the $Z_1 \log Z_1$-moment condition.

Lemma 10 can also be used to get very useful bounds for $\mathbf{P}(Z_n = k | Z_0 = \ell)$ which are not uniform in k. This will be achieved in the next lemma by specializing Lemma 10 to $h = 1$.

Lemma 12 *(Non-uniform bounds). There exist two positive constants A and δ such that*

$$c_n \mathbf{P}(Z_n = k | Z_0 = \ell) \leqslant A\, \mathrm{e}^{k/c_n} \ell^{-1/2}\, \mathrm{e}^{-\delta \ell}, \quad n, k \geqslant 1,\ \ell \geqslant \ell_0, \tag{101}$$

[*with ℓ_0 defined in* (87)].

Proof. By the branching property and the definition (85) of $S_\ell(h, n)$,

$$\mathbf{P}(Z_n = k | Z_0 = \ell) = \mathrm{e}^{kh/c_n} \big[f_n\big(\mathrm{e}^{-h/c_n}\big)\big]^\ell \mathbf{P}\big(S_\ell(h, n) = k\big). \tag{102}$$

Putting here $h = 1$ and multiplying both sides by c_n, we have

$$c_n \mathbf{P}(Z_n = k | Z_0 = \ell) \leqslant \mathrm{e}^{k/c_n} \big[f_n\big(\mathrm{e}^{-1/c_n}\big)\big]^\ell \max_{n,k \geqslant 1} c_n \mathbf{P}\big(S_\ell(1, n) = k\big). \tag{103}$$

Using Lemma 10 gives

$$c_n \mathbf{P}(Z_n = k | Z_0 = \ell) \leqslant A(1) \ell^{-1/2}\, \mathrm{e}^{k/c_n} \big[f_n\big(\mathrm{e}^{-1/c_n}\big)\big]^\ell. \tag{104}$$

From (77) the existence of a $\delta > 0$ follows such that $f_n(\mathrm{e}^{-1/c_n}) \leqslant \mathrm{e}^{-\delta}$ for all $n \geqslant 1$. Entering this into (104) finishes the proof. $\square$

2.3. On the limiting density function w

Recall from Section 1.1 that w denotes the density function of W, and $\psi = \psi_W$ its characteristic function.

Lemma 13 *(Bounds for the limiting density). There is a constant $A > 0$ such that*

$$w^{*\ell}(x) \leqslant A \left(\int_0^x w(t)\, \mathrm{d}t \right)^{\ell - \ell_0}, \quad x > 0,\ \ell \geqslant \ell_0. \tag{105}$$

Proof. Suppose $\alpha < \infty$, the case $\alpha = \infty$ can be treated similarly. By the inversion formula,

$$w^{*\ell_0}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}tx} \psi^{\ell_0}(t)\, \mathrm{d}t, \quad x > 0. \tag{106}$$

Hence,

$$A := \sup_{x > 0} w^{*\ell_0}(x) \leqslant \frac{1}{2\pi} \int_{-\infty}^{\infty} \big|\psi(t)\big|^{\ell_0} \mathrm{d}t. \tag{107}$$

We want to convince ourselves that $A < \infty$. It is well-known that ψ solves the equation

$$\psi(mu) = f\big(\psi(u)\big), \quad u \in \mathbb{R} \tag{108}$$

(e.g. [1, formula (6.1)]). Iterating, we obtain

$$\psi\big(m^\ell u\big) = f_\ell\big(\psi(u)\big), \quad u \in \mathbb{R},\ \ell \geqslant 1. \tag{109}$$

Thus, for $j \geqslant 0$,

$$\int_{m^j}^{m^{j+1}} \big|\psi(t)\big|^{\ell_0} \mathrm{d}t = m^j \int_1^m \big|\psi\big(tm^j\big)\big|^{\ell_0} \mathrm{d}t = m^j \int_1^m \big|f_j\big(\psi(t)\big)\big|^{\ell_0} \mathrm{d}t. \tag{110}$$

Since $W > 0$ has an absolute continuous law, $|\psi(t)| \leqslant C < 1$ for $t \in [1, m]$. Moreover, by (82), $|f_j(z)| \leqslant C p_1^j$ for z in a compact subset of D°. Therefore,

$$\int_{m^j}^{m^{j+1}} \left|\psi(t)\right|^{\ell_0} \mathrm{d}t \leqslant C m^j p_1^{j\ell_0} = C m^{j(1-\alpha\ell_0)} \tag{111}$$

by definition (8) of α. Consequently,

$$\int_1^\infty \left|\psi(t)\right|^{\ell_0} \mathrm{d}t \leqslant C \sum_{j=0}^\infty m^{j(1-\alpha\ell_0)} < \infty, \tag{112}$$

since $1 - \alpha\ell_0 < 0$. Analogously,

$$\int_{-\infty}^{-1} \left|\psi(t)\right|^{\ell_0} \mathrm{d}t < \infty. \tag{113}$$

Hence, A in (107) is finite. But $w^{*(\ell+1)}(x) = \int_0^x w^{*\ell}(x-y) w(y)\, \mathrm{d}y$, $x > 0$, and the claim follows by induction. □

2.4. A local central limit theorem

Recall the notation (85) for $S_\ell(h, n)$, $h \geqslant 0$, $\ell, n \geqslant 1$. By an abuse of notation, denote by $\psi_\ell = \psi_\ell^{h,n}$ the characteristic function of the random variable

$$\ell^{-1/2}\sigma^{-1}(h,n)\big(S_\ell(h,n) - \mathbf{E}S_\ell(h,n)\big), \tag{114}$$

where $\sigma(h, n) := \sqrt{\mathbf{E}(X_1(h,n) - \mathbf{E}X_1(h,n))^2}$. Note that by (86),

$$\psi_\ell^{h,n}(t) = \left(\mathrm{e}^{-\mathrm{i}t\ell^{-1/2}\sigma^{-1}(h,n)\mathbf{E}X_1(h,n)} \frac{f_n(\mathrm{e}^{-h/c_n + \mathrm{i}t\ell^{-1/2}\sigma^{-1}(h,n)})}{f_n(\mathrm{e}^{-h/c_n})}\right)^\ell. \tag{115}$$

Lemma 14 *(An Esseen type inequality). If $0 < h_1 \leqslant h_2 < \infty$, then there exist positive constants $C = C(h_1, h_2)$ and $\varepsilon = \varepsilon(h_1, h_2) < 1$ such that*

$$\sup_{h \in [h_1, h_2],\ n \geqslant 1} \left|\psi_\ell^{h,n}(t) - \mathrm{e}^{-t^2/2}\right| \leqslant C\ell^{-1/2}|t|^3 \mathrm{e}^{-t^2/3}, \quad |t| < \varepsilon\ell^{1/2},\ \ell \geqslant 1. \tag{116}$$

Proof. Put $\overline{X}_j(h,n) := X_j(h,n) - \mathbf{E}X_j(h,n)$. Using the global limit theorem from (3), one easily verifies that for some positive constants $C_1, \ldots, C_4$,

$$C_1 \leqslant \frac{\sigma(h,n)}{c_n} \leqslant C_2 \quad \text{uniformly in } h \in [h_1, h_2] \text{ and } n \geqslant 1 \tag{117}$$

and

$$C_3 \leqslant \frac{\mathbf{E}|\overline{X}_1(h,n)|^3}{c_n^3} \leqslant C_4 \quad \text{uniformly in } h \in [h_1, h_2] \text{ and } n \geqslant 1. \tag{118}$$

Consequently, the Lyapunov ratio $\mathbf{E}|\overline{X}_1(h,n)|^3/\sigma^3(h,n)$ is bounded away from zero and infinity. Applying now Lemma V.1 from [15] to the random variables $\overline{X}_1(h,n), \ldots, \overline{X}_\ell(h,n)$, we get the desired result. □

The next lemma is a key step in our development concerning the Böttcher case. Recall the notations $S_\ell := S_\ell(h,n)$ and $\sigma := \sigma(h,n)$ defined in (85) and after (114), respectively.

Lemma 15 *(Local central limit theorem). Suppose the offspring law is of type* (d,μ)*. If* $0<h_1\leqslant h_2<\infty$*, then*

$$\sup_{\substack{h\in[h_1,h_2]\\ n\geqslant 1}}\ \sup_{k:\ k\equiv\ell\mu^n(\operatorname{mod} d)}\left|\ell^{1/2}\sigma(h,n)\mathbf{P}\bigl(S_\ell(h,n)=k\bigr)-\frac{d}{\sqrt{2\pi}}\,\mathrm{e}^{-x_{k,\ell}^2(h,n)/2}\right|\underset{\ell\uparrow\infty}{\longrightarrow}0,$$

where $x_{k,\ell}:=x_{k,\ell}(h,n):=\ell^{-1/2}\sigma^{-1}(h,n)(k-\ell\mathbf{E}X_1(h,n))$.

Note that a local limit theorem, which would correspond to our case $h=0$ but concerning an offspring law with finite variance and with initial state tending to ∞, was derived by Höpfner [12, Theorem 1]. The following proof of our lemma is a bit simpler, since for $h>0$ the random variables $X_1(h,n)$ have finite moments of all orders (even if the underlying Z does not have finite variance).

Proof of Lemma 15. By (86) and the inversion formula,

$$\mathbf{P}(S_\ell=k)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{-\mathrm{i}tk}\left[\frac{f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t})}{f_n(\mathrm{e}^{-h/c_n})}\right]^\ell\mathrm{d}t. \tag{119}$$

Decomposing the unit circle,

$$\bigl\{\mathrm{e}^{\mathrm{i}t}\colon -\pi<t\leqslant\pi\bigr\}=\bigcup_{j=0}^{d-1}\bigl\{\varrho^j\,\mathrm{e}^{\mathrm{i}t}\colon -\pi d^{-1}<t\leqslant\pi d^{-1}\bigr\}, \tag{120}$$

where $\varrho:=\mathrm{e}^{2\pi\mathrm{i}/d}$, the latter integral equals

$$\sum_{j=0}^{d-1}\int_{-\pi d^{-1}}^{\pi d^{-1}}\varrho^{-jk}\mathrm{e}^{-\mathrm{i}tk}\left[\frac{f_n(\varrho^j\mathrm{e}^{-h/c_n+\mathrm{i}t})}{f_n(\mathrm{e}^{-h/c_n})}\right]^\ell\mathrm{d}t. \tag{121}$$

It is known (see, for instance, [1, p.105]) that for an offspring law of type (d,μ) we have

$$f_n\bigl(\varrho^j z\bigr)=\varrho^{j\mu^n}f_n(z),\quad n,\,j\geqslant 1,\ z\in D. \tag{122}$$

Therefore the latter sum equals

$$\int_{-\pi d^{-1}}^{\pi d^{-1}}\mathrm{e}^{-\mathrm{i}tk}\left[\frac{f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t})}{f_n(\mathrm{e}^{-h/c_n})}\right]^\ell\mathrm{d}t\sum_{j=0}^{d-1}\varrho^{-j(k-\ell\mu^n)}. \tag{123}$$

But $\varrho^{-j(k-\ell\mu^n)}\equiv 1$ for $k\equiv\ell\mu^n(\operatorname{mod} d)$. Altogether, for (119) we get

$$\mathbf{P}(S_\ell=k)=\frac{d}{2\pi}\int_{-\pi d^{-1}}^{\pi d^{-1}}\mathrm{e}^{-\mathrm{i}tk}\left[\frac{f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t})}{f_n(\mathrm{e}^{-h/c_n})}\right]^\ell\mathrm{d}t,\quad k\equiv\ell\mu^n(\operatorname{mod} d). \tag{124}$$

Using the substitution $t\to t/\ell^{1/2}\sigma$ and (115), we arrive at

$$\mathbf{P}(S_\ell=k)=\frac{d}{2\pi\,\ell^{1/2}\sigma}\int_{-\pi d^{-1}\ell^{1/2}\sigma}^{\pi d^{-1}\ell^{1/2}\sigma}\mathrm{e}^{-\mathrm{i}tx_{k,\ell}}\psi_\ell(t)\,\mathrm{d}t,\quad k\equiv\ell\mu^n(\operatorname{mod} d). \tag{125}$$

Fix $0<h_1\leqslant h_2<\infty$. Recall from (117) that

$$C_1\leqslant\inf_{h\in[h_1,h_2],\ n\geqslant 1}\frac{\sigma(h,n)}{c_n}\leqslant\sup_{h\in[h_1,h_2],\ n\geqslant 1}\frac{\sigma(h,n)}{c_n}\leqslant C_2 \tag{126}$$

for some $0<C_1<C_2$ (depending on h_1,h_2). Choose a positive

$$\varepsilon=\varepsilon(h_1,h_2)<C_1\pi d^{-1} \tag{127}$$

as in Lemma 14. Take any $A = A(h_1, h_2) > \varepsilon$ (to be specified later). Then the identity $\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}tx - t^2/2}\,\mathrm{d}t = \sqrt{2\pi}\,\mathrm{e}^{-x^2/2}$ and representation (125) imply that

$$\sup_{k:\ k \equiv \ell\mu^n (\mathrm{mod}\, d)} \left| \ell^{1/2}\sigma \mathbf{P}(S_\ell = k) - \frac{d}{\sqrt{2\pi}}\,\mathrm{e}^{-x_{k,\ell}^2/2} \right| \leqslant d(I_1 + I_2 + I_3 + I_4), \tag{128}$$

where

$$\begin{aligned} I_1 &:= \int_{-\varepsilon\ell^{1/2}}^{\varepsilon\ell^{1/2}} \left|\psi_\ell(t) - \mathrm{e}^{-t^2/2}\right| \mathrm{d}t, \qquad I_2 := \int_{|t|>\varepsilon\ell^{1/2}} \mathrm{e}^{-t^2/2}\,\mathrm{d}t, \\ I_3 &:= \int_{\varepsilon\ell^{1/2}<|t|<A\ell^{1/2}} \left|\psi_\ell(t)\right| \mathrm{d}t, \qquad I_4 := \int_{A\ell^{1/2}<|t|<\pi d^{-1}\ell^{1/2}\sigma} \left|\psi_\ell(t)\right| \mathrm{d}t. \end{aligned} \tag{129}$$

[Of course, I_4 disappears if $A(h_1, h_2) > \pi d^{-1}\sigma(h, n)$.]

Trivially, $I_2 \to 0$ as $\ell \uparrow \infty$. Further, due to Lemma 14, there is a $C = C(h_1, h_2)$ such that

$$I_1 \leqslant C\ell^{-1/2} \int_0^{\varepsilon\ell^{1/2}} t^3\,\mathrm{e}^{-t^2/3}\,\mathrm{d}t \leqslant C\ell^{-1/2} \underset{\ell\uparrow\infty}{\longrightarrow} 0. \tag{130}$$

Thus, it remains to show that the integrals I_3 and I_4 converge to zero as $\ell \uparrow \infty$, uniformly in the considered h and n.

First of all, using again (115) and substituting $t \to t\ell^{1/2}\sigma/c_n$, by (126) we obtain the following estimates

$$I_3 \leqslant C_2\ell^{1/2} \int_{\varepsilon/C_2<|t|<A/C_1} \left| \frac{f_n(\mathrm{e}^{-h/c_n + \mathrm{i}t/c_n})}{f_n(\mathrm{e}^{-h/c_n})} \right|^\ell \mathrm{d}t, \tag{131a}$$

$$I_4 \leqslant C_2\ell^{1/2} \int_{A/C_2<|t|<\pi d^{-1}c_n} \left| \frac{f_n(\mathrm{e}^{-h/c_n + \mathrm{i}t/c_n})}{f_n(\mathrm{e}^{-h/c_n})} \right|^\ell \mathrm{d}t. \tag{131b}$$

First we fix our attention on I_3. By (69),

$$f_n\big(\mathrm{e}^{-h/c_n + \mathrm{i}t/c_n}\big) \to \mathbf{E}\,\mathrm{e}^{-hW + \mathrm{i}tW} \quad \text{as } n \uparrow \infty, \tag{132}$$

uniformly in $h \in [0, h_2]$ and $t \in [0, A/C_1]$ [recall (127)]. It follows that

$$\frac{f_n(\mathrm{e}^{-h/c_n + \mathrm{i}t/c_n})}{f_n(\mathrm{e}^{-h/c_n})} \underset{n\uparrow\infty}{\longrightarrow} \frac{\mathbf{E}\,\mathrm{e}^{-hW + \mathrm{i}tW}}{\mathbf{E}\,\mathrm{e}^{-hW}} = \mathbf{E}\,\mathrm{e}^{\mathrm{i}tW(-h)}, \tag{133}$$

uniformly in $h \in [0, h_2]$ and $t \in [0, A/C_1]$ (with $W(-h)$ the Cramér transform of W). Since the $W(-h)$ have absolutely continuous laws, we have $|\mathbf{E}\,\mathrm{e}^{\mathrm{i}tW(-h)}| < 1$ for all $h \geqslant 0$ and $|t| > 0$. This inequality and continuity of $(h, t) \mapsto \mathbf{E}\,\mathrm{e}^{\mathrm{i}tW(-h)}$ imply that

$$\sup_{0\leqslant h\leqslant h_2,\ \varepsilon/C_2\leqslant|t|\leqslant A/C_1} \frac{|\mathbf{E}\,\mathrm{e}^{-hW + \mathrm{i}tW}|}{\mathbf{E}\mathrm{e}^{-hW}} < 1. \tag{134}$$

Using (133) and (134) we infer the existence of a positive constant $\eta = \eta(h_1, h_2) < 1$ and an $n_1 = n_1(h_1, h_2) \geqslant 1$ such that for $n \geqslant n_1$,

$$\sup_{0\leqslant h\leqslant h_2,\ \varepsilon/C_2\leqslant|t|\leqslant A/C_1} \left| \frac{f_n(\mathrm{e}^{-h/c_n + \mathrm{i}t/c_n})}{f_n(\mathrm{e}^{-h/c_n})} \right| \leqslant \eta. \tag{135}$$

Applying (135) to the bound of I_3 in (131a), we conclude that

$$I_3 \leqslant CA\ell^{1/2}\eta^\ell \to 0 \quad \text{as } \ell \uparrow \infty, \tag{136}$$

uniformly in $h \in [h_1, h_2]$ and $n \geqslant n_1$. (The remaining n will be considered below.)

Next, we prepare for the estimation of I_4. Since $f_n(\mathrm{e}^{-h/c_n}) \geqslant f_n(\mathrm{e}^{-h_2/c_n})$ for $0 \leqslant h \leqslant h_2$, and $f_n(\mathrm{e}^{-h_2/c_n}) \to \mathbf{E}\,\mathrm{e}^{-h_2 W} > 0$ as $n \uparrow \infty$ [recall (132)], there is a positive constant $C = C(h_2)$ such that

$$\left|\frac{f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t})}{f_n(\mathrm{e}^{-h/c_n})}\right| \leqslant C\left|f_n\big(\mathrm{e}^{-h/c_n+\mathrm{i}t}\big)\right| \tag{137}$$

for all $t \in \mathbb{R}$, $0 \leqslant h \leqslant h_2$, and $n \geqslant 1$.

At this point we have to distinguish between the Schröder and Böttcher cases. Actually, we proceed with the Böttcher case $\alpha = \infty$, which is the only case we need later, and leave the other case for the reader. Applying the second case of (78) to (137), we obtain the estimate

$$\left|\frac{f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t})}{f_n(\mathrm{e}^{-h/c_n})}\right| \leqslant C\exp\big[-\mu^{n-j+1}\log\theta^{-1}\big], \tag{138}$$

$0 \leqslant h \leqslant h_2$, $t \in J_j$, and $1 \leqslant j \leqslant n$. Since $\mu \geqslant 2$, the right-hand side of (138) is bounded by

$$C\exp\big[-\mu^{n-j}\log\theta^{-1}\big]\exp\big[-\mu^{n-j}\log\theta^{-1}\big]. \tag{139}$$

Evidently, there exists an $n_2 = n_2(h_2)$ such that $C\exp[-\mu^{n-j}\log\theta^{-1}] \leqslant 1$ for $1 \leqslant j \leqslant n - n_2$. Therefore,

$$\left|\frac{f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t})}{f_n(\mathrm{e}^{-h/c_n})}\right| \leqslant \exp\big[-\mu^{n-j}\log\theta^{-1}\big], \tag{140}$$

if $0 \leqslant h \leqslant h_2$, $t \in J_j$, and $1 \leqslant j \leqslant n - n_2$. But $|J_j| \leqslant 2c_{j-1}^{-1}\pi d^{-1}$, hence

$$\int_{J_j}\left|\frac{f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t})}{f_n(\mathrm{e}^{-h/c_n})}\right|^{\ell}\mathrm{d}t \leqslant 2c_{j-1}^{-1}\pi d^{-1}\exp\big[-\ell\mu^{n-j}\log\theta^{-1}\big]. \tag{141}$$

Summing over the considered j gives

$$\int\limits_{c_{n-n_2}^{-1}\pi d^{-1} \leqslant |t| \leqslant \pi d^{-1}}\left|\frac{f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t})}{f_n(\mathrm{e}^{-h/c_n})}\right|^{\ell}\mathrm{d}t \leqslant 2\pi d^{-1}\sum_{j=1}^{n-n_2} c_{j-1}^{-1}\exp\big[-\ell\mu^{n-j}\log\theta^{-1}\big],$$

$0 \leqslant h \leqslant h_2$ and $n \geqslant n_2$. Substituting $t \to t/c_n$ and using (94), we arrive at

$$\begin{aligned}\int\limits_{\pi d^{-1}m^{n_2} \leqslant |t| \leqslant \pi d^{-1}c_n}\left|\frac{f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t/c_n})}{f_n(\mathrm{e}^{-h/c_n})}\right|^{\ell}\mathrm{d}t &\leqslant 2\pi d^{-1}\sum_{j=1}^{n-n_2} m^{n-j+1}\exp\big[-\ell\mu^{n-j}\log\theta^{-1}\big]\\ &\leqslant 2\pi d^{-1}\sum_{j=1}^{\infty} m^{j+1}\exp\big[-\ell\mu^{j}\log\theta^{-1}\big] \leqslant C\mathrm{e}^{-C'\ell}\end{aligned} \tag{142}$$

with constants C, C', uniformly in $h \in [h_1, h_2]$ and $n \geqslant n_2$. Choosing now A so large that $\pi d^{-1}m^{n_2} \leqslant A/C_2$, we conclude from (131b) that

$$I_4 \leqslant C\ell^{1/2}\,\mathrm{e}^{-C'\ell} \to 0 \quad \text{as } \ell \uparrow \infty, \tag{143}$$

uniformly in $h \in [h_1, h_2]$ and $n \geqslant n_2$.

Finally, we consider all $n \leqslant n^* := n_1 \vee n_2$. By definition, as in (73),

$$\frac{f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t/c_n})}{f_n(\mathrm{e}^{-h/c_n})} = \sum_{j=0}^{\infty}\mathbf{P}\big(X_1(h,n) = \mu^n + jd\big)\,\mathrm{e}^{(\mathrm{i}t/c_n)(\mu^n+jd)}. \tag{144}$$

Hence, since the set $\{\mathrm{e}^{-\mathrm{i}t/c_n}\colon\ t \in [\varepsilon/C_2, \pi d^{-1}c_n]\}$ does not contain the dth roots of unity,

$$\sup_{t \in [\varepsilon/C_2, \pi d^{-1}c_n]}\left|\frac{f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t/c_n})}{f_n(\mathrm{e}^{-h/c_n})}\right| =: \theta_n(h) \ < \ 1. \tag{145}$$

From the continuity $(h,t) \to f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t/c_n})$ it follows that the function θ_n is continuous, too. Therefore,

$$\sup_{h\in[h_1,h_2]} \theta_n(h) =: \bar{\theta}_n < 1. \tag{146}$$

Combining (145) and (146),

$$\max_{n\leqslant n^*} \sup_{\substack{h\in[h_1,h_2]\\ t\in[\varepsilon/C_2,\pi d^{-1}c_n]}} \left| \frac{f_n(\mathrm{e}^{-h/c_n+\mathrm{i}t/c_n})}{f_n(\mathrm{e}^{-h/c_n})} \right| \leqslant \bar{\theta} \tag{147}$$

for some $\bar{\theta} < 1$. Substituting this into (131) gives

$$I_3 + I_4 \leqslant C\ell^{1/2}\bar{\theta}^{\ell} \to 0 \quad \text{as } \ell \uparrow \infty, \tag{148}$$

and the proof is finished. □

3. Proof of the main results

3.1. Schröder case (proof of Theorem 4)

Let f, k, and a_k be as in Theorem 4. Recall that $p_0 = 0$ by our convention. By the Markov property,

$$\mathbf{P}(Z_{a_k+j} = k) = \sum_{\ell=1}^{\infty} \mathbf{P}(Z_j = \ell)\mathbf{P}(Z_{a_k} = k | Z_0 = \ell) \tag{149}$$

and

$$\mathbf{P}(Z_{a_k+j} \leqslant k) = \sum_{\ell=1}^{\infty} \mathbf{P}(Z_j = \ell)\mathbf{P}(Z_{a_k} \leqslant k | Z_0 = \ell). \tag{150}$$

Step 1° (*Proof of* (42)). Using Lemma 12, we get for $N \geqslant \ell_0$ the estimate

$$c_{a_k} \sum_{\ell=N}^{\infty} \mathbf{P}(Z_j = \ell)\mathbf{P}(Z_{a_k} = k | Z_0 = \ell) \leqslant C \frac{\mathrm{e}^{k/c_{a_k}}}{N^{1/2}} f_j\left(\mathrm{e}^{-\delta}\right) \tag{151}$$

for some constant $\delta > 0$. By (4a), and since $c_{a_k-1} < k \leqslant c_{a_k}$ by the definition of a_k,

$$m^{-1} \leqslant \frac{c_{a_k-1}}{c_{a_k}} \leqslant \frac{k}{c_{a_k}} \leqslant 1. \tag{152}$$

On the other hand, by (82),

$$f_j\left(\mathrm{e}^{-\delta}\right) \leqslant C p_1^j. \tag{153}$$

Thus, from (151),

$$p_1^{-j} c_{a_k} \sum_{\ell=N}^{\infty} \mathbf{P}(Z_j = \ell)\mathbf{P}(Z_{a_k} = k | Z_0 = \ell) \leqslant \frac{C}{N^{1/2}}. \tag{154}$$

By [8, Lemma 9],

$$\lim_{n\uparrow\infty} \frac{1}{2\pi} \int_{-\pi d^{-1}c_n}^{\pi d^{-1}c_n} f_n^{\ell}\left(\mathrm{e}^{\mathrm{i}t/c_n}\right) \mathrm{e}^{-\mathrm{i}tx}\, \mathrm{d}t = w^{*\ell}(x) \tag{155}$$

uniformly in $x \in [m^{-1}, 1]$. This together with

$$c_{a_k}\mathbf{P}(Z_{a_k} = k | Z_0 = \ell) = \frac{d}{2\pi} \int_{-\pi d^{-1}c_{a_k}}^{\pi d^{-1}c_{a_k}} f_{a_k}^{\ell}\left(\mathrm{e}^{\mathrm{i}t/c_{a_k}}\right) \mathrm{e}^{-\mathrm{i}tk/c_{a_k}}\, \mathrm{d}t, \quad \ell \equiv k \ (\mathrm{mod}\ d), \tag{156}$$

(see [1, p. 105]) and (152) gives

$$\sup_{k\geqslant\tilde{k}:\ k\equiv\ell(\operatorname{mod}d)} \left|c_{a_k}\mathbf{P}(Z_{a_k}=k|Z_0=\ell)-dw^{*\ell}(k/c_{a_k})\right| \underset{\tilde{k}\uparrow\infty}{\longrightarrow} 0. \tag{157}$$

Since $k\equiv 1(\operatorname{mod}d)$, the previous statement holds for all $\ell\equiv 1(\operatorname{mod}d)$. For other ℓ, the probabilities $\mathbf{P}(Z_j=\ell)$ disappear. Thus, by (157),

$$\sum_{\ell=1}^{N-1}\mathbf{P}(Z_j=\ell)\mathbf{P}(Z_{a_k}=k|Z_0=\ell)=dc_{a_k}^{-1}\left[\sum_{\ell=1}^{N-1}\mathbf{P}(Z_j=\ell)w^{*\ell}(k/c_{a_k})\right](1+\epsilon_{N,k}), \tag{158}$$

where $\epsilon_{N,k}\in\mathbb{R}$ satisfies $\sup_{k\geqslant\tilde{k}}|\epsilon_{N,k}|\to 0$ as $\tilde{k}\uparrow\infty$, for each fixed N. Further, using Lemma 13, one can easily verify that there exist two constants C and $\eta\in(0,1)$ such that $w^{*\ell}(k/c_{a_k})\leqslant C\eta^{\ell}$ for all $\ell\geqslant 1$ and k. Thus,

$$\sum_{\ell=N}^{\infty}\mathbf{P}(Z_j=\ell)w^{*\ell}(k/c_{a_k})\leqslant C\sum_{\ell=N}^{\infty}\mathbf{P}(Z_j=\ell)\eta^{\ell}. \tag{159}$$

But for every $\eta_1\in(\eta,1)$,

$$\sum_{\ell=N}^{\infty}\mathbf{P}(Z_j=\ell)\eta^{\ell}\leqslant\left(\frac{\eta}{\eta_1}\right)^N f_j(\eta_1)\leqslant C\left(\frac{\eta}{\eta_1}\right)^N p_1^j, \tag{160}$$

where in the last step we used (82). Inequalities (159) and (160) imply

$$\sum_{\ell=N}^{\infty}\mathbf{P}(Z_j=\ell)w^{*\ell}(k/c_{a_k})\leqslant Cp_1^j\,\mathrm{e}^{-\delta N}, \tag{161}$$

for all j,k,N and some constant $\delta>0$. Combining (149), (158), (154) and (161), we have

$$\mathbf{P}(Z_{a_k+j}=k)=dc_{a_k}^{-1}\left[\sum_{\ell=1}^{\infty}\mathbf{P}(Z_j=\ell)w^{*\ell}(k/c_{a_k})\right](1+\epsilon_{N,k})+\mathrm{O}\big(c_{a_k}^{-1}p_1^jN^{-1/2}\big), \tag{162}$$

where the O-term applies to $j,k,N\uparrow\infty$. By (109),

$$m^{-j}w(x/m^j)=\sum_{\ell=1}^{\infty}\mathbf{P}(Z_j=\ell)w^{*\ell}(x),\qquad j\geqslant 0,\ x>0. \tag{163}$$

Substituting this into (162), we arrive at

$$\mathbf{P}(Z_{a_k+j}=k)=dc_{a_k}^{-1}m^{-j}w\big(km^{-j}/c_{a_k}\big)(1+\epsilon_{N,k})+\mathrm{O}\big(c_{a_k}^{-1}p_1^jN^{-1/2}\big).$$

By (29), (152), and the definition (8) of α,

$$dc_{a_k}^{-1}m^{-j}w\big(km^{-j}/c_{a_k}\big)\geqslant Cc_{a_k}^{-1}m^{-\alpha j}=Cc_{a_k}^{-1}p_1^j,\quad \text{for all } k. \tag{164}$$

Therefore,

$$\mathbf{P}(Z_{a_k+j}=k)=dc_{a_k}^{-1}m^{-j}w\big(km^{-j}/c_{a_k}\big)\big(1+\epsilon_{N,k}+\mathrm{O}\big(N^{-1/2}\big)\big), \tag{165}$$

where the O-term now applies to $N\uparrow\infty$, uniformly for all k,j. Letting first $\tilde{k}\uparrow\infty$ and then $N\uparrow\infty$, we see that (42) is true.

Step 2° (*Proof of* (43)). Trivially, for independent and identically distributed non-negative random variables $X_1,\dots,X_n$ we have

$$\mathbf{P}(X_1+\cdots+X_n<x)\leqslant\mathbf{P}\Big(\max_j X_j<x\Big)=\big[\mathbf{P}(X_1<x)\big]^n,\quad x\geqslant 0. \tag{166}$$

Hence,

$$\mathbf{P}(Z_{a_k}\leqslant k|Z_0=\ell)\leqslant\big[\mathbf{P}(Z_{a_k}\leqslant k)\big]^{\ell}. \tag{167}$$

Further, from (152) and (3),

$$\mathbf{P}(Z_{a_k} \leqslant k) \leqslant \mathbf{P}\big(c_{a_k}^{-1} Z_{a_k} \leqslant 1\big) \underset{k\uparrow\infty}{\longrightarrow} \int_0^1 w(x)\,\mathrm{d}x. \tag{168}$$

Therefore, since $w > 0$ on all of $(0, \infty)$, there exists an $\eta \in (0, 1)$ such that $\mathbf{P}(Z_{a_k} \leqslant k) \leqslant \eta$ for all n large enough. Thus,

$$\sum_{\ell=N}^{\infty} \mathbf{P}(Z_j = \ell)\mathbf{P}(Z_{a_k} \leqslant k | Z_0 = \ell) \leqslant \sum_{\ell=N}^{\infty} \mathbf{P}(Z_j = \ell)\eta^{\ell} \tag{169}$$

for all N sufficiently large. Taking into account (160), we conclude that

$$\sum_{\ell=N}^{\infty} \mathbf{P}(Z_j = \ell)\mathbf{P}(Z_{a_k} \leqslant k | Z_0 = \ell) \leqslant C p_1^j\,\mathrm{e}^{-\delta N} \tag{170}$$

for N sufficiently large and some $\delta > 0$. By the same arguments,

$$\sum_{\ell=N}^{\infty} \mathbf{P}(Z_j = \ell) F^{*\ell}(k/c_{a_k}) \leqslant C p_1^j\,\mathrm{e}^{-\delta N}, \tag{171}$$

where $F(x) := \mathbf{P}(W < x),\ x \geqslant 0$, and $F^{*\ell}$ is the ℓ-fold convolution.

On the other hand, the continuity of F and (3) yield that $\mathbf{P}(Z_{a_k} \leqslant c_{a_k} x | Z_0 = \ell) \to F^{*\ell}(x)$ uniformly in $x \geqslant 0$. Therefore,

$$\lim_{k\uparrow\infty} \sup_{i\geqslant 1} \big|\mathbf{P}(Z_{a_k} \leqslant i | Z_0 = \ell) - F^{*\ell}(i/c_{a_k})\big| = 0. \tag{172}$$

Combining (150), (170), (171), and (172), we arrive at

$$\mathbf{P}(Z_{a_k+j} \leqslant k) = \left[\sum_{\ell=1}^{\infty} \mathbf{P}(Z_j = \ell) F^{*\ell}(k/c_{a_k})\right](1 + \epsilon_{N,k}) + \mathrm{O}\big(p_1^j\,\mathrm{e}^{-\delta N}\big) \tag{173}$$

with the same meaning of $\epsilon_{N,k}$ and the O-term as in the previous step of proof. Since $\mathbf{P}(Z_j = 1) = p_1^j$ and $F(k/c_{a_k}) \geqslant F(m^{-1}) > 0$ by (152), we obtain

$$p_1^j\,\mathrm{e}^{-\delta N} \leqslant C\,\mathrm{e}^{-\delta N} \sum_{\ell=1}^{\infty} \mathbf{P}(Z_j = \ell) F^{*\ell}(k/c_{a_k}). \tag{174}$$

Combining this inequality with (173) gives

$$\mathbf{P}(Z_{a_k+j} \leqslant k) = \left[\sum_{\ell=1}^{\infty} \mathbf{P}(Z_j = \ell) F^{*\ell}(k/c_{a_k})\right]\big(1 + \epsilon_{N,k} + \mathrm{O}\big(\mathrm{e}^{-\delta N}\big)\big). \tag{175}$$

Integrating both parts of (163), one has

$$F\big(y/m^k\big) = \sum_{\ell=1}^{\infty} \mathbf{P}(Z_k = \ell) F^{*\ell}(y), \quad k \geqslant 0,\ y > 0. \tag{176}$$

Thus,

$$\mathbf{P}(Z_{a_k+j} \leqslant k) = F\left(\frac{k}{c_{a_k} m^j}\right)\big(1 + \epsilon_{N,k} + \mathrm{O}\big(\mathrm{e}^{-\delta N}\big)\big). \tag{177}$$

Letting again first $\tilde{k} \uparrow \infty$ and then $N \uparrow \infty$ finishes the proof. $\square$

Remark 16 *(Proof in the case $p_0 > 0$).* We indicate now how to proceed with the proof of Theorem 4 in the remaining case $p_0 > 0$. Here in the representation (149) one has additionally to take into account that

$$\mathbf{P}(Z_{a_k} = k | Z_0 = \ell) = \sum_{j=1}^{\ell} \binom{\ell}{j} f_{a_k}^{\ell-j}(0)\big(1 - f_{a_k}(0)\big)^j \mathbf{P}\left\{\sum_{i=1}^{j} Z_{a_k}^{(i)} = k \big| Z_{a_k}^{(i)} > 0,\ 1 \leqslant i \leqslant j\right\}, \tag{178}$$

where the $Z^{(1)}, Z^{(2)}, \ldots$ are independent copies of Z. Then instead of Lemma 12 we need

$$c_n \mathbf{P}\left\{\sum_{i=1}^{j} Z_n^{(i)} = k \big| Z_n^{(i)} > 0,\ 1 \leqslant i \leqslant j\right\} \leqslant A\, \mathrm{e}^{k/c_n} j^{-1/2}\, \mathrm{e}^{-\delta j}, \quad n, k \geqslant 1,\ j \geqslant \ell_0.$$

But this is valid by

$$\mathbf{E}\big\{z^{Z_n^{(1)}} \big| Z_n^{(1)} > 0\big\} = \frac{f_n(z) - f_n(0)}{1 - f_n(0)} \underset{n\uparrow\infty}{\longrightarrow} \frac{\mathsf{S}(z) - \mathsf{S}(0)}{1 - q}, \tag{179}$$

uniformly in z from compact subsets of D°. This indeed follows from (9).

3.2. Böttcher case (proof of Theorem 6)

From the Markov property,

$$\mathbf{P}(Z_n = k_n) = \sum_{\ell=\mu^{n-b_n}}^{\infty} \mathbf{P}(Z_{n-b_n} = \ell)\mathbf{P}(Z_{b_n} = k | Z_0 = \ell). \tag{180}$$

Using (102) and Lemma 10, we obtain the following estimate

$$c_{b_n}\mathbf{P}(Z_{b_n} = k_n | Z_0 = \ell) \leqslant A(h)\ell^{-1/2}\big[\mathrm{e}^{hk_n/\ell c_{b_n}} f_{b_n}\big(\mathrm{e}^{-h/c_{b_n}}\big)\big]^{\ell}. \tag{181}$$

From the definition of b_n it immediately follows that

$$2k_n \leqslant c_{b_n}\mu^{n-b_n} = c_{b_n-1}\mu^{n-b_n+1}\left(\frac{c_{b_n}}{\mu c_{b_n-1}}\right) \leqslant 2k_n\frac{m}{\mu}. \tag{182}$$

Hence,

$$\frac{hk_n}{\ell c_{b_n}} \leqslant \frac{h}{2} \tag{183}$$

for $\ell \geqslant \mu^{n-b_n}$. Therefore,

$$c_{b_n}\mathbf{P}(Z_{b_n} = k_n | Z_0 = \ell) \leqslant A(h)\ell^{-1/2}\big[\mathrm{e}^{h/2} f_{b_n}\big(\mathrm{e}^{-h/c_{b_n}}\big)\big]^{\ell}. \tag{184}$$

It is known (see, for example, [1], Corollary III.5.7), that $\mathbf{E}W = 1$ if $\mathbf{E}Z_1 \log Z_1 < \infty$ and $\mathbf{E}W = \infty$ otherwise. This means, that for the Laplace transform $\varphi = \varphi_W$ of W we have $\mathrm{e}^{h/2}\varphi(h) < 1$ for all small enough h. Thus, due to the global limit theorem (3), there exist $\delta < 1$ and $h_0 > 0$ such that $\mathrm{e}^{h_0/2} f_n(\mathrm{e}^{-h_0/c_n}) \leqslant \mathrm{e}^{-\delta}$ for all large enough n. Hence,

$$c_{b_n}\mathbf{P}(Z_{b_n} = k_n | Z_0 = \ell) \leqslant A\ell^{-1/2}\,\mathrm{e}^{-\delta\ell}. \tag{185}$$

Inserting (185) into (180), we obtain

$$c_{b_n}\mathbf{P}(Z_n = k_n) \leqslant A\mu^{-(n-b_n)/2} f_{n-b_n}\big(\mathrm{e}^{-\delta}\big), \tag{186}$$

consequently,

$$\mu^{b_n-n}\log\big[c_n\mathbf{P}(Z_n = k_n)\big] \leqslant \mu^{b_n-n}C + \mu^{b_n-n}\log\left(\frac{c_n}{c_{b_n}}\right) + \frac{\log f_n(\mathrm{e}^{-\delta})}{\mu^{n-b_n}}. \tag{187}$$

Since $c_n/c_{b_n} \leqslant m^{n-b_n}$ and $\mu^{n-b_n} = m^{\beta(n-b_n)}$, $\mu^{b_n-n}\log(c_n/c_{b_n}) \to 0$ as $n \uparrow \infty$. Thus,

$$\limsup_{n\uparrow\infty} \mu^{b_n-n}\log\big[c_n\mathbf{P}(Z_n = k_n)\big] \leqslant \limsup_{n\uparrow\infty} \frac{\log f_{n-b_n}(\mathrm{e}^{-\delta})}{\mu^{n-b_n}}. \tag{188}$$

Using (21), we arrive at the desired upper bound.

We show now that (62b) holds for $\log \mathbf{P}(Z_n \leqslant k_n)$. First of all we note that for arbitrary non-negative random variable X and all $x, h \geqslant 0$,

$$\mathbf{P}(X \leqslant x) \leqslant e^{hx}\mathbf{E}\, e^{-hX}. \tag{189}$$

Applying this bound to the process Z starting from ℓ individuals and taking into account (183), we have

$$\mathbf{P}(Z_{b_n} \leqslant k_n | Z_0 = \ell) \leqslant \left[e^{hk_n/\ell c_{b_n}} f_{b_n}\left(e^{-h/c_{b_n}}\right)\right]^\ell \leqslant \left[e^{h/2} f_{b_n}\left(e^{-h/c_{b_n}}\right)\right]^\ell. \tag{190}$$

As we argued in the derivation of (185), this gives

$$\mathbf{P}(Z_{b_n} \leqslant k_n | Z_0 = \ell) \leqslant e^{-\delta\ell}. \tag{191}$$

Consequently, by the Markov property,

$$\mathbf{P}(Z_n \leqslant k_n) \leqslant f_{n-b_n}\left(e^{-\delta}\right). \tag{192}$$

Taking logarithms and using (21), we obtain (62b).

Let us verify the lower bounds in Theorem 6. By (180),

$$\mathbf{P}(Z_n = k_n) \geqslant \mathbf{P}\left(Z_{n-b_n} = \mu^{n-b_n}\right)\mathbf{P}\left(Z_{b_n} = k_n | Z_0 = \mu^{n-b_n}\right). \tag{193}$$

From (102),

$$\mathbf{P}\left(Z_{b_n} = k_n | Z_0 = \mu^{n-b_n}\right) > \left[f_{b_n}\left(e^{-h/c_{b_n}}\right)\right]^{\ell_n}\mathbf{P}\left(S_{\ell_n}(h, b_n) = k_n\right), \tag{194}$$

where $\ell_n = \mu^{n-b_n}$.

Consider the equation

$$c_{b_n}^{-1}\mathbf{E}X_1(h, b_n) = \frac{f'_{b_n}(e^{-h/c_{b_n}})\, e^{-h/c_{b_n}}}{c_{b_n} f_{b_n}(e^{-h/c_{b_n}})} = x. \tag{195}$$

Evidently,

$$\left.\frac{f'_{b_n}(e^{-h/c_{b_n}})\, e^{-h/c_{b_n}}}{f_{b_n}(e^{-h/c_{b_n}})}\right|_{h=0} = m^{b_n} \tag{196}$$

and

$$\left.\frac{f'_{b_n}(e^{-h/c_{b_n}})\, e^{-h/c_{b_n}}}{f_{b_n}(e^{-h/c_{b_n}})}\right|_{h=\infty} = \mu^{b_n}. \tag{197}$$

From these identities and monotonicity of $f'_{b_n}(e^{-h/c_{b_n}})\, e^{-h/c_{b_n}}/f_{b_n}(e^{-h/c_{b_n}})$ it follows that (195) has a unique solution $h_n(x)$ for $\mu^{b_n}c_{b_n}^{-1} < x < m^{b_n}c_{b_n}^{-1}$. Analogously one shows that the equation $\varphi'(h)/\varphi(h) = -x$ has also a unique solution $h(x)$. By the integral limit theorem (3), the right-hand side in (195) converges to $-\varphi'(h)/\varphi(h)$ and consequently, $h_n(x) \to h(x)$ as $n \uparrow \infty$. Further, by (182),

$$\frac{\mu}{2m} \leqslant x_n := \frac{k_n}{c_{b_n}\ell_n} \leqslant \frac{1}{2}. \tag{198}$$

Thus,

$$h(\mu/2m) \leqslant \liminf_{n\uparrow\infty} h_n \leqslant \limsup_{n\uparrow\infty} h_n \leqslant h\left(\frac{1}{2}\right), \tag{199}$$

where $h_n := h_n(x_n)$. This means that there exist h_* and h^* such that $h_* \leqslant h_n \leqslant h^*$ for all $n \geqslant 1$. From the definition of h_n and from (195), it immediately follows that $\mathbf{E}S_{\ell_n}(h_n, b_n) = k_n$. Thus, applying Lemma 15, we get

$$\lim_{n\uparrow\infty}\left|\ell_n^{1/2}\sigma(h_n, b_n)\mathbf{P}\left(S_{\ell_n}(h_n, b_n) = k_n\right) - \frac{d}{\sqrt{2\pi}}\right| = 0. \tag{200}$$

Recall that by (117) we have $\sigma(h_n, b_n) \geqslant C c_{b_n}$. Hence,

$$\liminf_{n\uparrow\infty} \ell_n^{1/2} c_{b_n} \mathbf{P}\big(S_{\ell_n}(h_n, b_n) = k_n\big) \geqslant C > 0. \tag{201}$$

Moreover, since $f_{b_n}(\mathrm{e}^{-h_n/c_{b_n}}) \geqslant f_{b_n}(\mathrm{e}^{-h^*/c_{b_n}})$ and $f_j(\mathrm{e}^{-h^*/c_j}) \to \mathbf{E}\,\mathrm{e}^{-h^* W} > 0$, there exists a $\theta > 0$ such that

$$f_{b_n}\big(\mathrm{e}^{-h/c_{b_n}}\big) \geqslant \theta \tag{202}$$

for all n. Applying these bounds to the right-hand side in (194), we find that

$$\liminf_{n\uparrow\infty} \mu^{b_n-n} \log\big[c_n \mathbf{P}\big(Z_{b_n} = k_n | Z_0 = \mu^{n-b_n}\big)\big] \geqslant -C. \tag{203}$$

Using this inequality and (21) to bound the right-hand side in (193), we conclude that

$$\liminf_{n\uparrow\infty} \mu^{b_n-n} \log\big[c_n \mathbf{P}(Z_n = k_n)\big] \geqslant -C, \tag{204}$$

i.e. (62a) is proved.

Next we want to extend this result to $\mathbf{P}(Z_n \leqslant k_n)$. Obviously,

$$\mathbf{P}(Z_n \leqslant k_n) \geqslant \mathbf{P}(Z_{n-b_n} = \ell_n)\mathbf{P}(Z_{b_n} \leqslant k_n | Z_0 = \ell_n). \tag{205}$$

Then, using (102) with $h = h_n$, we have

$$\mathbf{P}(Z_n \leqslant k_n) \geqslant \mathbf{P}(Z_{n-b_n} = \ell_n)\big[f_n\big(\mathrm{e}^{-h_n/c_{b_n}}\big)\big]^{\ell_n} \mathbf{P}\big(S_{\ell_n}(h, b_n) \leqslant k_n\big). \tag{206}$$

By the central limit theorem,

$$\lim_{n\uparrow\infty} \mathbf{P}\big(S_{\ell_n}(h, b_n) \leqslant k_n\big) = \frac{1}{2}. \tag{207}$$

From this statement and (202) it follows that

$$\liminf_{n\uparrow\infty} \mu^{b_n-n} \log \mathbf{P}(Z_n \leqslant k_n) \geqslant \liminf_{n\uparrow\infty} \mu^{b_n-n} \log \mathbf{P}\big(Z_{n-b_n} = \mu^{n-b_n}\big) + \log\theta. \tag{208}$$

Recalling (17), the proof of Theorem 6 is complete. □

Acknowledgements

We are grateful to Peter Mörters and Vladimir Vatutin for a discussion of our original preprint [10]. We also thank anonymous referees for a careful reading and their suggestions that led to an improvement of the exposition.

References

[1] A. Asmussen, H. Hering, Branching Processes, Progress in Probab. and Statistics, vol. 3, Birkhäuser Boston Inc., Boston, MA, 1983.
[2] K.B. Athreya, P.E. Ney, The local limit theorem and some related aspects of supercritical branching processes, Trans. Amer. Math. Soc. 152 (2) (1970) 233–251.
[3] J.D. Biggins, N.H. Bingham, Near-constancy phenomena in branching processes, Math. Proc. Cambridge Philos. Soc. 110 (3) (1991) 545–558.
[4] J.D. Biggins, N.H. Bingham, Large deviations in the supercritical branching process, Adv. Appl. Probab. 25 (4) (1993) 757–772.
[5] N.H. Bingham, On the limit of a supercritical branching process, J. Appl. Probab. 25A (1988) 215–228.
[6] S. Dubuc, La densite de la loi-limite d'un processus en cascade expansif, Z. Wahrsch. Verw. Gebiete 19 (1971) 281–290.
[7] S. Dubuc, Etude theorique et numerique de la fonction de Karlin–McGregor, J. Analyse Math. 42 (1982) 15–37.
[8] S. Dubuc, E. Seneta, The local limit theorem for the Galton–Watson process, Ann. Probab. 4 (1976) 490–496.
[9] W. Feller, An Introduction to Probability Theory and its Applications, vol. II, second ed., John Wiley and Sons, New York, 1971.
[10] K. Fleischmann, V. Wachtel, Lower deviation probabilities for supercritical Galton–Watson processes, WIAS Berlin, Preprint No. 1025 of April 28, 2005, http://www.wias-berlin.de/publications/preprints/1025/.
[11] K. Fleischmann, V. Wachtel, Large deviations for sums defined on a Galton–Watson process, WIAS Berlin, Preprint No. 1135 of May 29, 2006.
[12] R. Höpfner, Local limit theorems for non-critical Galton–Watson processes with or without immigration, J. Appl. Probab. 19 (1982) 262–271.
[13] M. Kuczma, Functional Equations in a Single Variable, PWN, Warszaw, 1968.
[14] P.E. Ney, A.N. Vidyashankar, Local limit theory and large deviations for supercritical branching processes, Ann. Appl. Probab. 14 (2004) 1135–1166.
[15] V.V. Petrov, Sums of Independent Random Variables, Springer-Verlag, Berlin, 1975.
[16] E. Seneta, Regularly Varying Functions, Lecture Notes in Math., vol. 508, Springer-Verlag, Berlin, 1976.